DOCUMENT RESUME

ED 151 598 EC 103 914

AUTHOR Budoff, Milton
TITLE Learning Potential Assessment of Educable Mentally Retarded Students. Final Report.
INSTITUTION Research Inst. for Educational Problems, Cambridge, Mass.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO OHDS-77-31104
PUB DATE May 76
GRANT NE-G-3-0016
NOTE 477p.; For related information, see EC 103 856

EDRS PRICE MF-$0.83 HC-$26.11 Plus Postage.
DESCRIPTORS Cultural Differences; *Educable Mentally Handicapped; Elementary Secondary Education; Evaluation Methods; Exceptional Child Research; Mentally Handicapped; Research; *Student Evaluation; *Test Bias; *Test Validity
IDENTIFIERS *Learning Potential Assessment

ABSTRACT

The final report of a project on learning potential assessment for educable mentally retarded students reviews previous research in the area, presents 13 formal reports completed under the grant, and summarizes conclusions and implications on the topic. The first section explains the differences between learning potential (LP) assessment and traditional IQ evaluation, and reviews research on LP measurement (including the Kohs Block Design LP procedure, and the series Learning Potential Test), and validity of LP assessment (including reviews of LP and paired associate learning and the relationship of LP to personality variables). Thirteen reports contained in the second section address such topics as the following: the effectiveness of training on the Raven LP measure scores of children from diverse racial and socioeconomic backgrounds, effectiveness of LP procedure in improving retarded and nonretarded children's problem solving skills, and validity of LP assessment. Implications of such a training-based assessment approach are discussed. Included in five appendixes are the manual, directions, and scoring key for the Raven Learning Potential Test and a discussion of the Raven Learning Potential procedure for the moderately and severely retarded. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED151998

FINAL REPORT

United States
Department of Health, Education, and Welfare
National Institute of Education

Grant NE-G-00-3-0016

Learning Potential Assessment of Educable Mentally Retarded Students

Milton Budoff, Ph.D.

Research Institute for Educational Problems
29 Ware Street
Cambridge, Massachusetts 02138

May, 1976

EC 103914

TABLE OF CONTENTS

Prefatory Information

Foreword

Staff List

List of Publications

Abstract

Page

PART I. Introduction and Overview of Past Work 1

Measurement of Learning Potential 5

A. Kohs Block Design Learning Potential Procedure 7

B. The Development of a Learning Potential Measure Based on Raven Progressive Matrices 11

C. The Series Learning Potential Test 17

Validity of Learning Potential Assessment 25

A. The Problem of the Criterion of Learning Potential 25

B. Social, Demographic, and Psychometric Correlates 28

C. Educational Implication of Learning Potential Status 33

D. Learning Potential and Paired Associate Learning 35

E. Learning Potential in Relation to Personality Variables 36

F. Summary of Interview with Special and Regular Class Adolescents 49

G. Learning Potential and Mother's Perceptions of her Child's Placement in an EMR Class 55

H. Learning Potential and Post School Adjustment 56

Summary of the Construct Validity for Learning Potential Assessment 59

PART II. Studies Performed on this Project

Study 1

Corman, L. & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure. RIEP-Print No. 56, 1973. [ED 086 924]

Study 2

Budoff, M. & Corman, L. The effectiveness of a group training procedure on the Raven learning potential measure with children from diverse racial and socioeconomic backgrounds. RIEP-Print No. 58, 1973. [ED 086 923]

Study 3

Corman, L. & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure with black and white special class students. RIEP-Print No. 57, 1973. [EC 060 957, ED 085 969]

Study 4

Corman, L. & Budoff, M. IQ and learning potential measurements of general intelligence: A comparison of relationships. In D. A. A. Primrose (Ed.), Proceedings of the Third Congress of the International Association for the Scientific Study of Mental Deficiency (Vol. 1). Stirlingshire, Scotland: IASSMD, 1975.

Study 5

Corman, L. & Budoff, M. Factorial composition of Raven's Progressive Matrices in the context of learning potential assessment. RIEP-Print No. 62, 1973 /1975.

Study 6

Budoff, M. & Corman, L. Factor analysis of error types on Raven Progressive Matrices: A failure to obtain logical structure. RIEP-Print No. 79, 1975.

Study 7

Budoff, M. & Corman, L. Effectiveness of a learning potential procedure in improving problem-solving skills of retarded and nonretarded children. RIEP-Print No. 63, 1975.

Study 8

Corman, L. & Budoff, M. The picture word game: A nonverbal test of the ability to use language-related symbols. *RIEP-Print* No. 77, 1974. [ED 108 439]

Study 9

Budoff, M. A further study of the educational validity of learning potential assessment: Response to objectives-based curricula. 1976.

Study 10

Corman, L, & Budoff, M. Factor structures of Spanish-speaking and non-Spanish-speaking children on Raven's Progressive Matrices. *Journal of Educational and Psychological Measurement*, 1974, *34*, 977-981.

Study 11

Budoff, M., Corman, L., & Gimon, A. An educational test of learning potential assessment with Spanish-speaking youth. *Interamerican Journal of Psychology*, 1976, *10*, 13-24. (ED 108 436]

Study 12

Budoff, M.,& Hamilton, J.L. Optimizing test performance of the moderately and severely mentally retarded. *American Journal of Mental Deficiency*, 1976, *81*, 49 - 57.

Study 13

Babad, E. & Bashi, J. Final Report: An educational test of the validity of learning potential measurement. Ford Foundation, June 1975. *RIEP-Print* No. 91, 1975.

	Pages
PART III. Summary and Implications	1 - 8

APPENDICES

Appendix A	Manual for the Raven Learning Potential Test Directions for Administration of Raven Matrices Scoring Key, Raven Progressive Matrices Answer Sheets, Raven Progressive Matrices
Appendix B	Sample Short Form of Raven
Appendix C	The Raven Instructions in Spanish
Appendix D	Raven Learning Potential Procedure for the Moderately and Severely Retarded
Appendix E	Picture Word Game Instructions Picture Word Game Training Book Picture Word Game

Foreword

This final report on Grant NE-G-00-3-0016, "Learning Potential Assessment of Educable Mentally Retarded Students" from the National Institute of Education reports the results of further studies of learning potential (LP) measurement of intelligence, as contrasted with measures generated by traditional intelligence tests. The original proposal which was submitted was conceived as a partial followup to a program project grant previously funded by the Bureau of Education for the Handicapped which covered a broad range of research problems. In the grant renewal, as submitted, studies in two areas addressed in the prior project were proposed: continuing the line of research in learning potential assessment procedures and their validity, and a controlled longitudinal study of the effects of mainstreaming EMR children who were then attending special classes. This latter study was cast in an aptitude-by-treatment interaction conceptual framework, in which LP status was one aptitude variable, and the treatments were special and regular class placement. It was a larger scale expansion over a longer time frame of a small scale study by Budoff and Gottlieb (1976).

The final award, however, was for less than 20% of the requested amount for about half the time requested. The study of the effects of mainstreaming was specifically excluded. No specific work plan was developed consequent to the awarding except that the work be confined to the area of learning potential measurement. This final report contains formal presentations of the results relating to problems described in Studies 1 through 4 of the section concerned with learning potential assessment (pages 18-26 of proposal). Due to the constraints imposed by the limited funding, the curriculum study described in the proposal was abandoned, and this study (no.5) was coordinated with an extant experimental curriculum development project conducted by the Bureau for Children with Retarded Mental Development (BCRMD) of the New York City Board of Education. In this study, LP and IQ were used as predictor variables to describe the progress made by special class children on an objectives-based curriculum.

In addition to the studies proposed, prior work with Spanish-speaking children was completed with the aid of support from this grant award. The results of two studies are presented, a factor analysis of the Raven Progressive Matrices with Spanish-speaking children, and a study validating the learning potential procedure using an educational intervention. Earlier studies had shown in a more limited manner that these procedures constituted a fruitful direction.

A further study was completed which continued to validate the utility of this approach for moderately and severely mentally retarded persons. Utilizing several testing formats that incorporated training on nonverbal tasks of perception and reasoning, some moderately and severely mentally retarded children following training solved more difficult problems than was expected based on their IQs and history of prior accomplishments.

The last study describes the results of a cross-cultural test of the applicability of the learning potential approach to children from homes one generation removed from a non-Western background. Through a cooperative working relationship with Professors Elisha Babad and Joseph Bashi of the Hebrew University of Jerusalem, one curriculum-type validity study was completed with Israeli Jewish children, and is included in this report. This latter study was partially supported, only to the extent that the experimental design and the materials were developed in consultation between the principal investigator and the Israeli investigators.

In sum, this final report will first present an introduction and overview of the prior work in the area of learning potential by the Institute and other investigators. Section 2 will present the individual formal reports of the studies partly or wholly completed under the NIE grant. The final section will present some conclusions and implications of this work.

The Appendices present a manual which includes the instructions for administering the Raven Matrices training procedure employed in the studies described with American EMR and normal children and the actual test forms and the training booklet (Appendix A). More recent work subsequent to completion of this project indicates that training can be conducted more effectively with these groups when training booklets are not used. By focusing the active attention of the students on the blackboard through the use of the slides and dispensing with the training booklets, we have some initial evidence that the children attend more to the training procedure. The booklets appear to be distracting and pose many problems in drawing. With many children, especially younger ones, the shift in attention from the slide presentation to the booklets and back again, was overly distracting, especially when drawing in the appropriate design was often difficult for them. A translation of the original procedure was employed with the Puerto Rican children. Appendix B consists of a sample short form of 6 items used to help children who seem to require some advance orientation to the requirements of the Raven Test answer sheet. Appendix C presents the instructions in Spanish, although they represent an earlier version of the training procedure. Appendix D presents the instructions for the training for moderately and severely mentally retarded students. Appendix E consists of the instructions, training, and test materials for the Picture-Word Game.

PROFESSIONAL STAFF

Elisha Y. Babad, Ph.D.	Duke University Psychology	Research Associate at RIEP, presently Hebrew University, Jerusalem
Milton Budoff, Ph.D.	University of Chicago Human Development	Director, RIEP Developmental and Cognitive Psychology
Louise Corman, Ph.D	Boston College Educational Research	Statistician and Programmer Test Construction, Experimental Design
Alex Gimon, Ph.D.	Yeshiva University Research in Mental Retardation	Research Associate Bilingual Assessment and Education
Jay Gottlieb, Ph.D.	Yeshiva University Research in Mental Retardation	Assistant Director, RIEP Social Psychology of MR, Special Ed., and Educational Evaluation
James L. Hamilton, Ph.D.	University of Missouri Columbia, Special Education, MR	Research Associate Special Education of Moderate and Severely Retarded Persons
Robert H. Harrison, Ph.D.	Pennsylvania State Clinical Psychology	Consultant Statistical Analysis and Research Design
Gary N. Siperstein, Ph.D.	Yeshiva University Psychological Research in MR	Research Associate Experimental Psychology and Ed. Evaluation
Joan Lerner Stein, Ed.D.	Harvard University Laboratory of Human Development	Research Associate Developmental and Cognitive Psychology
Cydney Stoler, M.Ed.	Northeastern University Special Education, Mental Retardation	Research Associate Development of Training Programs for Severely Retarded Children

List of Publications

Budoff, M. Corman, L. & Gimon, A. An educational test of learning potential assessment with Spanish-speaking youth. *Interamerican Journal of Psychology*, 1976, *10*, 13-24. [ED 108 436]

Budoff, M. & Hamilton, J. L. Optimizing test performance of the moderately and severely mentally retarded. *American Journal of Mental Deficiency*, 1976, *81*, 49-57.

Corman, L. & Budoff, M. Factor structures of Spanish-speaking and non-Spanish-speaking children on Raven's Progressive Matrices. *Journal of Educational and Psychological Measurement*, 1974, *34*, 977-981.

Corman, L. & Budoff, M. IQ and learning potential measurements of general intelligence: A comparison of relationships. In D. A. A. Primrose (Ed.), *Proceedings of the Third Congress of the International Association for the Scientific Study of Mental Deficiency* (Vol. I), Stirlingshire, Scotland: IASSMD, 1975.

Project Title: Learning Potential Assessment of Educable Mentally Retarded Students

Abstract

Learning Potential assessment measures the ability to reason in a test-train-test paradigm among IQ-defined educable mental retardates (EMRs). The measure appears to be culture-fair to low income and minority group children. The test-train-test paradigm provides practice and instruction on problem-relevant strategies in a competence inducing context. Intelligence is defined as ability to profit from instruction and apply it in nontrained instances of the problem. Among IQ-defined low income EMRs, high able (LP) students have been shown to be educationally handicapped while low improving students tend to function as mentally retarded students.

The research contined to refine the learning potential assessment procedures structured around Raven Progressive Matrices by testing a completely group administered procedure, developing a language analogue measure of learning potential. The procedure was tested for cultural bias by race (black and white), income level (middle and low), and regular and special class children. Direct tests of the validity of learning potential assessment were conducted by testing the ability to learn using curricula specially designed to tap the cognitive strengths of these children.

PART I

Introduction and Overview of Past Work

Pages

Measurement of Learning Potential 5 - 7

- A. Kohs Block Design Learning Potential Procedure 7 - 11
- B. The Development of a Learning Potential Measure Based on Raven Progressive Matrices 11 - 16
- C. The Series Learning Potential Test 17 - 24

Validity of Learning Potential Assessment

- A. The Problem of the Criterion of Learning Potential 25 - 28
- B. Social, Demographic, and Psychometric Correlates 28 - 32
- C. Educational Implications of Learning Potential Status 33 - 35
- D. Learning Potential and Paired Associate Learning 35 - 36
- E. Learning Potential in Relation to Personality Variables 36 - 49
- F. Summary of Interview with Special and Regular Class Adolescents 49 - 55
- G. Learning Potential and Mother's Perceptions of Her Child's Placement in an EMR Class 55 - 56
- H. Learning Potential and Post School Adjustment 56 - 58

Summary of the Construct Validity for Learning Potential Assessment 59 - 60

Introduction and Overview of Past Work:

Learning potential assessment represents a response to the growing dissatisfaction with the validity of traditional IQ scores for poor, non-middle-class, and frequently, non-white population. IQ tests (group and individual) primarily assess the extent to which a child has spontaneously acquired knowledge and school-related proficiencies from his natural environment. The IQ test is based on the assumption that children have equal opportunities of access to school-preparatory experiences, and, thus, that differences in acquisition of knowledge, information, and skills reflect differences in (inborn) intellectual abilities. Unfortunately, poor and/or non-white children often lack access to appropriate school-preparatory experiences in early childhood, and perform poorly on IQ tests administered during the school years, indicating a low probability for success in academic school programs. If the IQ scores were viewed as predictors of scholastic success, these children's inferior IQs should indicate the need for altered school programs to forestall a predicted failure. Rather, the low IQ scores are usually interpreted to indicate lower (inborn)intelligence, and result in the misclassification for educational purposes of substantial proportions of non-middle-class children. It is underestimating their capabilities since their prior experiences have not prepared them for working with and solving the types of problems presented.

IQ testing is a particularly discriminatory practice when used as the primary criterion for classifying children as mentally retarded and segregating them into special classes, or into any other low educational track that may restrict the child's future educational opportunities. The most glaring evidence of misclassification is the gross over-representation of poor and/or non-white children in the special classes for the educable mentally retarded (EMR). More than 85% of all children in these classes are poor and/or non-white; the middle class children invariably have organic brain complications or severe emotional disturbance. Clearly a new approach to the testing of mental ability is required which would minimize the cultural biases of the tests, and offer insights into potential abilities.

The Research Institute for Educational Problems has been studying learning potential assessment as an alternative procedure for estimating general ability to reason. The learning potential (LP) assessment approach is based on a conceptualization of intelligence which stresses trainability, or the ability to profit from learning experiences. It assumes that the prior and continuing experiences of poor and/or non-white children do not allow them to spontaneously acquire school-relevant skills in their natural environment. The learning potential paradigm provides subjects with training experiences directly relevant to the reasoning task presented, so they can apply, in microcosm, their problem-solving ability and show whether they can improve their performance on the task. The improved performance indicates problem-solving capability not evident when training is not provided as part of the test administration. Learning potential assessment replaces the traditional product-oriented test with a three-stage procedure which includes a pretest, a training session, and a posttest.

The early work at RIEP was concerned with studying the correlates of response following training. Three patterns of response are evident when the subjects are drawn from the narrow Binet type IQ range of 60 to 80 IQ. A proportion of the subjects (high scorers) perform exceedingly well on the nonverbal pretest, solving problems at levels attained by higher IQ children of their age; another group of children (gainers) solved problems more successfully following the training procedure; while the third group (nongainers) essentially did not profit from the training experience.

The mass of data collected with the learning potential instruments during these development years has focused on understanding the differences in response by students in special classes and institutions for the mentally retarded to the learning potential assessment. The hypothesis was that those who profited from training and performed more effectively on the nonverbal reasoning task are educationally handicapped but not mentally retarded. That is, the concern originally, and in a continuing way, has been with the false positive problem which has resulted in a substantial incidence of misclassification of low income children as educable mental retardates, particularly those from non-white and non-English-speaking backgrounds.

The data we have accumulated indicate a broad range of consistency and support for the differences shown by psychometrically defined educable mental retardates (EMRs) in response to the test-train-retest paradigm which constitutes the learning potential assessment strategy. More able children, by the learning potential criterion, indicate areas of more adequate functioning when performance does not depend on satisfactory progress in school. Those children who have demonstrated their competence on the LP task (e.g., the Kohs Block Designs) also tended to score higher on other nonverbal or minimally verbal tasks, such as the Wechsler Intelligence Scale for Children Performance Subtest IQ (PIQ), and Raven Progressive Matrices. The personality data, whether behavioral or verbal, suggests a greater sense of personal adequacy in the more able children, greater flexibility on a concept shift task (Budoff & Pagell, 1968), more flexibility and ability to delay response (Mankinen, 1970b), ability to manage frustrations more positively (Budoff & Pines, 1970), a view of oneself as less neurotic or maladjusted (Harrison & Budoff, 1972), the tendency to set aspiration levels consonent with one's performance even under the stress of failure (Harrison, Singer, Budoff, & Folman, 1972; Mankinen, 1970a) and the feeling that one is not manifestly different socially from one's "nonretarded" peers (Folman &

Budoff, 1972a). However, they do realistically feel inadequate in academic areas (Folman & Budoff, 1972). The students who seem to be most able by the learning potential criterion are also those who seem to be making more concrete (low level) and realistic vocational plans (Folman & Budoff, 1971).

The competence and understanding of the more able children can more generally be demonstrated by their differential ability to build empirically meaningful relationships among objects and/or events, when they are not required to state verbally the abstract elements embodied in their empirical understanding. The more able (LP) students' ability to perform in this manner was clearly demonstrated in a study in which they were taught electricity empirically by manipulating batteries, bulbs, and wires. Following teaching they were able to display understanding of the principles of electricity on a nonverbal task to the same degree as low achieving higher IQ children who were taught the curriculum. In some instances, they performed as well as average achieving children (Budoff, Meskin, & Harrison, 1971). The performance of special class children seems to indicate their difficulties are in the verbal expression of their understanding (Budoff, Meskin, & Harrison, 1971; Harrison, Greenberg, & Budoff, 1975; Suzman, 1973; McCaffrey, 1973). For example, when asked to give verbal responses for their correct empirical responses on a minimally verbal electricity test, the special class students gave few causal explanations and mainly provided more or less vague descriptions of the electrical problems. The regular class students, by contrast, gave a higher proportion of causal explanations. Suzman (1973) who allowed EMRs to paraphrase embedded sentences verbally or in mime, also reported that the difficulty of special class children appears to be in production of verbal responses, not in the understanding of the meaning of the sentence. The differences in the verbal paraphrases of embedded sentences by regular and special class students were not apparent when the paraphrase was performed in mime, with few utterances required. However, these deficiencies may be remediable. Mandelkorn and Corman (1973) have shown that the more able students, by the LP criterion, profited from a short unit which taught verbal communication skills to a greater degree than the less able special class students.

The most tenable hypothesis by which to explain these consistent differences between more and less able special class students is to consider the high able group by the learning potential criterion (high scorers and gainers) as educationally handicapped. These are children who have been unable to progress satisfactorily in school for whatever reasons, but who seem to have ability not usually ascribed to them.

Those who do not profit from the experience of training on the nonverbal reasoning tasks (nongainers) seem to function on a wide variety of tasks as mentally retarded children are described. The high able group by the LP criterion may be those children who have been designated by some as "pseudo-feebleminded." These are school-aged children diagnosed as mentally retarded. As adults in the wider opportunity system of the society, they become economically and socially independent, confounding the previous diagnosis which guided the educational opportunities made available to them. As a result, these children leave school as functionally illiterate. For a large proportion of EMR children (about 2/3 of those noninstitutionalized and over 15 years) the LP procedure evoked a positive response and indications of previously unsuspected ability to reason. This proportion, coincidentally, is the approximate proportion of school-aged "EMRs" who attain an independent social and economic adjustment during their adult years.

This evidence of ability to learn empirically should remove the more able students by the LP criterion from consideration as mentally retarded, though they seem to suffer from serious verbal-expressive difficulties which hamper their educational progress. In at least one instance, these verbal problems may be amenable to training (Mandelkorn & Corman, 1973).

The work with learning potential assessment started with a concern for developing a more equitable means of discriminating ability among severely low achieving school children who had been classified as educable mentally retarded. More recently, we have become concerned with using the learning potential assessment paradigm with children from backgrounds of high educational risk. These children, though not officially labeled mentally retarded, also experience considerable difficulty in school and are generally viewed as having limited scholastic aptitude (IQs below 100) for whom the tested low IQ score signals the need for revised programs, not a degradation of the curricula offerings as customarily has occurred in special class settings. In a general way, we shall demonstrate in this report that a training-based assessment procedure is applicable to a broad range of persons for whom the assumption of equal access to experiences common to middle class Western persons are not met. Thus, providing training minimizes the effects of inappropriate prior experience with reasoning problems of the type commonly employed on tests, and provides a more optimal estimate of a person's general ability to reason. The evidence from our work suggests strongly that with low income and/or non-white or non-English-speaking or non-Western students, a training-based assessment process represents a more culture-fair means of measuring general intelligence when intelligence is defined as the ability to learn and profit from experience.

The next section reviews the procedures by which learning potential can be assessed as of the start of this project. Three training procedures were available: the block designs found in the Kohs Block Designs Test, the problems presented in the Raven Progressive Matrices, and a new measure, wholly developed at RIEP, the Series Learning Potential Test (SLPT). Subsequent sections present the work completed in this project, and the final section draws some implications of this training-based assessment approach for the measurement of a person's general ability to profit from his experiences, i.e., his intelligence. This section will relate this work to the recent arguments of Jensen and Mercer. Jensen (1969) argues that Level II tasks, e.g., the reasoning tasks found on the Raven Progressive Matrices distinguish the brighter from dumber persons who can not be trained. Utilizing training appropriate to the Raven Progressive Matrices, as in the learning potential assessment procedure, a substantial proportion of IQ-defined EMRs can function at the level of their average IQ chronologically-aged peers. Mercer's argument for pluralistic culturally adjusted norms does demonstrate the increased capability of the child from low income, and often culturally different backgrounds. However, it leaves us at the same deadend as does the traditional IQ measurement. By employing assessment procedures in which training appropriate to the criterion tasks are routinely employed, one can not only demonstrate increased capability but one can also graphically illustrate the child's ability to learn, to improve his capability. If training-based assessment is routinely employed in individual or small group contexts, the tester or teacher can also describe the qualitative features of the child's learning --that is, his natural learning styles and approaches, what impediments he exhibits, how (easily) one can help him work through these obstacles during acquisition of his understanding of the problems, and his attempts to apply these understandings in appropriate circumstances. Thus, the distinctive contribution of training-based assessment to understanding a child's strengths and difficulties may suggest a more truly culture-fair approach to assessing aptitudes for school, for vocational opportunities, etc., in which the effects of prior experience on the outcome which prejudices the child's or adult's scores in the traditional testing format based on measuring cumulative acquisitions is at least partially minimized.

The Measurement of Learning Potential:

This section will describe the learning potential instruments available prior to the start of this project. This section describes the test tasks, the principles of training, the procedure, and some characteristics of the learning potential instrument. The instructions for each instrument are available upon request as part of a specimen set.

The LP assessment approach is based on a conceptualization of intelligence which stresses trainability, or the ability to profit from learning experiences. It assumes that the prior and continuing experiences of poor and/or non-white children do not allow them to spontaneously acquire school-relevant skills in their natural environment. The learning potential paradigm provides subjects with training experiences directly relevant to the reasoning task presented, so they can apply, in microcosm, their problem-solving ability and show whether they can improve their performance on the task. The improved performance indicates problem-solving capability not evident when training is not provided as part of the test administration. In the most usual format, learning potential assessment replaces the traditional product-oriented one-session test with a three-stage procedure which includes a pretest, training session(s), and a posttest.

As will be evident in this report, we have begun experiments with other formats. For the requirements of developing a nonverbal assessment of the child's ability to handle a language-like task, we had to teach the children the novel "language" before we could test the extent to which they could apply it. (See Corman & Budoff, 1975 in this report.) Also, because of the exigencies of the particular population, i.e., moderately and severely mentally retarded persons, we developed a paradigm in which training is provided as part of the testing procedure. This approach has proved useful with this population (see Hamilton & Budoff, 1974) and may also be relevant in applications to preschool and kindergarten children.

Prior to the start of this project, three measures of learning potential existed. The original measure made use of training on tasks similar to the Kohs Block Designs. Most of the construct validity data available on learning potential has utilized this procedure. This measure is most useful with children older than 11 years because the complexity of visual organizational skills required are too complex prior to that age.

A second procedure involved training on problems similar to those on the Raven Progressive Matrices. This procedure was developed to serve as a complementary measure to the Kohs Learning Potential Test, since Sets C, D, and E of the Raven test overlap the age range appropriate to the Kohs task. The easier sets (A, AB, B) permitted work with younger children as early as kindergarten. Raven Progressive Matrices is a standardized, nonverbal intelligence test which measures conceptual thinking with four types of problems: discrete pattern completions, continuity and reconstruction of simple and complex structures, reasoning by analogy in a 2 X 2 format, and linoleum block completions (Corman & Budoff, 1973). Group and individual training procedures exist for the entire age range. This procedure can be used with very low IQ or illiterate adults as well as children in kindergarten and first grade.

The first entirely group-administered learning potential measure was the Series Learning Potential Test (SLPT) which was wholly developed at the Research Institute for Educational Problems by Dr. Elisha Babad, now at the Hebrew University School of Education (Jerusalem, Israel). The Series Test is also a nonverbal reasoning task. Each item presents a series of stimuli, meaningful pictures, or geometric figures, and a blank cell. The subjects must identify among the multiple choices the one which best completes the series. The concepts in a series of pictures vary systematically by type of figure, size (large/small), color (black/white), or orientation (up/down, or left/right). Thus, a blank cell may require a large/black puppy dog. The child is trained to identify each concept separately, trace its pattern of occurrence in the series by "singing the tune" and cross out the choices that are inappropriate for a correct solution. As will be indicated below, this test has de facto ceiling problems and discriminates effectively in the span from 6 to 8 years of age. Dr. Babad is continuing to work with it in Israel with Israeli Jewish and Arab populations. We have set it aside and have developed a group training version of the Raven LP procedure for class-size administrations.

During 1972-1973, learning potential assessment was adapted for use with moderately and severely mentally handicapped children. These children are especially penalized in traditional testing situations, which demand responses not in their repertoire. Their variability of response and their poor initial performance dictate assessment strategies that allow the child several trials by which to attain more optimal determination of his capability on a task. One initial study (Hamilton, 1972) found promising correlations between the post-training scores of severely retarded adolescents and other psychometric and teacher rating measures. Further work in this area is described in Study 12 in this report.

A. The Kohs Block Design Learning Potential Procedure

Budoff's (1967, 1968, 1969, Budoff & Friedman, 1964) first attempt to measure learning potential used Kohs' (1923) original Block Design Test. Using enlarged designs drawn to the scale of one-inch cubes for testing and training, he devised a training program which taught strategies useful for solving block design problems. This particular task was selected because (1) it does not involve verbalization or verbal exchanges with the tester, (2) is not previously associated with school types of work and the subjects have not experienced failure with the task, (3) the task involves abstract reasoning which is represented by a few basic strategies which can be taught. With these strategies, a capable student can then solve quite difficult, sixteen block problems (4 X 4 designs).

1. Description of test and training materials.

The materials consist of a test series of 15 of the original Kohs Block Designs, including five designs with sixteen blocks. The designs are arranged in order of increasing difficulty. Design 7 from the Wechsler Adult Intelligence Scale was added. The designs are printed on 5 x 6 inch white cards to double the scale of the original Kohs designs so that the stimulus designs and the block constructions were equivalent in size. A design of four blocks is drawn as a two-inch square; one of nine blocks as a three-inch square; one of 16 blocks as a four-inch square. This modification was suggested by Goldstein and Sheerer's (1941) adaptation as a means of simplifying the task for the S. The four colors, i.e., red, white, blue, and yellow, are retained. The usual one-inch cubes were used as the blocks. The five coaching designs consist of three four-block designs (C from WISC, 3 and 7 from the Kohs series) and two nine-block designs (5 from the WISC and 8 from the WAIS). The coaching designs are printed in the same format and dimensions as the test designs.

2. Procedure.

The 16 test designs are administered individually three times: prior to coaching, and then one day, and again one month following coaching. Kohs' instructions for administering the block design test are used at each session. A sample problem is demonstrated by the examiner. The child has to construct it correctly before the remainder of items are presented. Testing is discontinued after three successive failures.

The training procedure emphasizes the following principles: (1) the subject's probability of success on a coaching item is maximized by the possibility of working down to the simplest elements in the design. Praise and encouragement are freely given. (2) The principle of motorically checking his construction, block by block, against the design card is built into the coaching procedure. It was assumed that the retardate is distractable and displays poor work habits. He is required, therefore, to actively point, block by block, to his constructed design and the corresponding blocks on the design cards. This is done on all the stimulus cards in which all the blocks on a stimulus card are delineated. It was hoped thereby to encourage a more planned and systematic work approach and to allow the subject to see concretely the success he is achieving. (3) The concept of two-color blocks was emphasized repeatedly as the elements composing the design, e.g., the process of building a stripe was demonstrated.

Training for an item was so designed that the subject has to solve the problem initially from a stimulus card in which the blocks are undifferentiated. If he fails to solve the problem, he is presented one row of the design at a time. Initially the blocks composing the row are not outlined. On succeeding presentations, if the subject fails to align the blocks correctly, the design is progressively outlined, e.g., see progression Figure 1. On the stimulus cards in which all the blocks were completely outlined (B3, C2, F3 and on the final card (H), the subject is required to check whether each block in his design corresponds exactly with the comparable block on the stimulus card. When two rows of the design are completed, the experimenter demonstrates how they fit together to form the corresponding portion of the pictured design (Figure 1d and 1g). After the entire design is learned in this piecemeal fashion, the subject is asked to reconstruct the pattern from the final stimulus card in the coaching series in which all the blocks were outlined. The coaching series for the four-block items was simpler, though based on the same principles.

(See Figure 1, printed on next page)

3. Results.

With adolescent educable mentally retarded subjects, some trained subjects showed substantial increases from pretest to posttest. Post-posttests with no intervening training, showed no decreases to initial level, while some Ss showed slight and continuing increases in competence. Training did not uniformaly increase all scores. Rather, there were three patterns of response to training: subjects who demonstrated little or no gain following coaching (nongainers), subjects who demonstrated quite marked gains (gainers), and a third group who did not need the coaching, as they scored highly on the pretest, demonstrating initial ability which could not be enhanced by the further training because of ceiling effects (high scorers).

Learning potential was conceived as a procedure which sought to maximize understanding of the task and to indicate potential not indicated in the low IQ or poor school achievement record. Learning potential status was defined

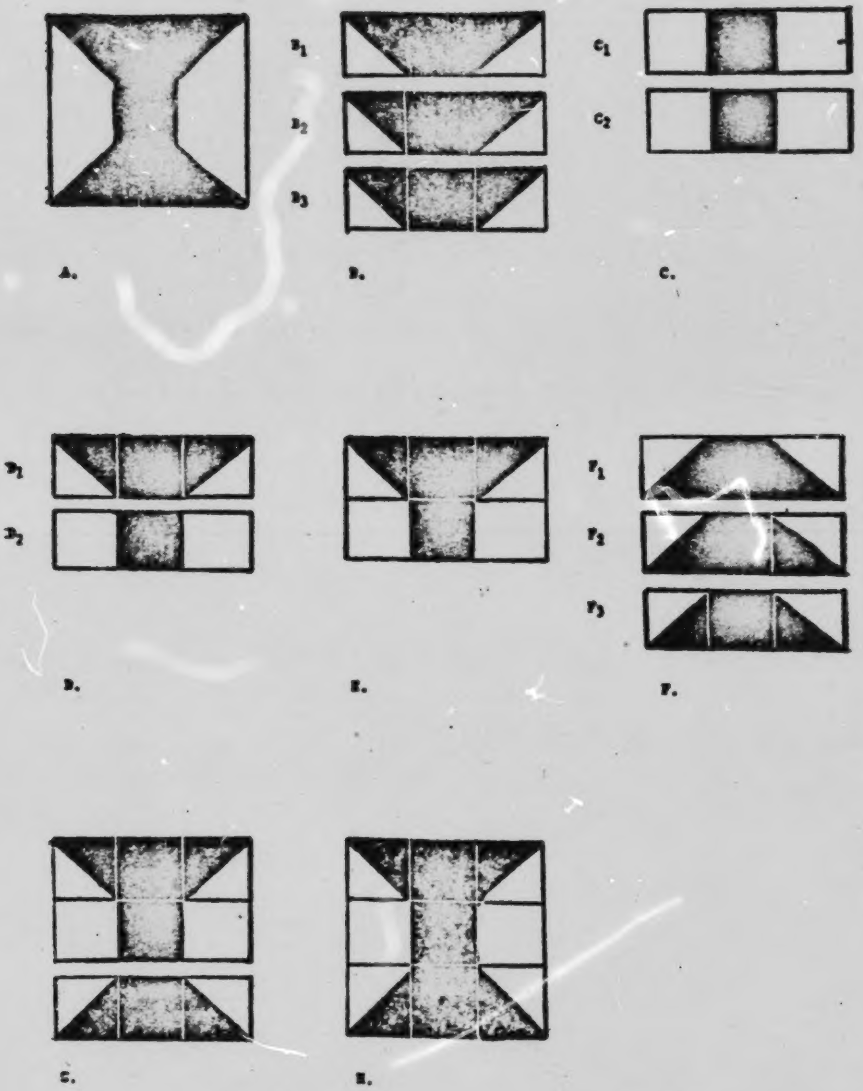


FIG. 1. Coaching procedure illustrated for nine block design (#4).
1-A. Initial design card.
1-B. Stimulus cards for top row of the design.
1-C. Stimulus cards for middle row of the design.
1-D. Card showing relationship of upper two rows of the design.
1-E. Card showing the two rows placed together.
1-F. Stimulus cards for lowest row of the design.
1-G. Card showing the relationship of the upper two rows and the lowest row of the design.
1-H. Final card with all blocks outlined.

conceptually as improvement in performance following training (gain score). Operationally, gainers were those individuals who improved their performance following training by four or more designs. Four designs were selected as the crucial point, as the mean practice gain of non-coached subjects was 1.2. To minimize the false positive problem, a cutoff point three times this increase was selected. Nongainers improved either posttraining score by less than four designs. Particularly with community special class students, some students performed very well prior to training on the non-verbal task. These high scorers were defined as individuals who solved one or more difficult 3 x 3 or 4 x 4 designs on the pretest. Substantial proportions of subjects fell into these learning potential categories.

The initial studies sought to describe the correlates of response to training. The data validating learning potential assessment for IQ-defined EMRs is largely drawn from work with the Kohs learning potential task. These data will be presented in the section on validity of learning potential assessment.

4. Reliability and stability of Kohs learning potential test.

The Kohs learning potential procedure results in a stable assessment of ability to reason. Correlations obtained between immediate and delayed (one month) posttraining scores range from .89 with large diverse samples, to .953 between immediate and delayed retest (one month interval) and .866 six months after coaching with smaller institutionalized samples (Budoff & Friedman, 1964; Budoff, 1967). Correlations of .701 and .835 between pretest and immediate posttest scores were computed from the scores of non-trained samples of EMR institutionalized and community special class adolescents.

B. The Development of a Learning Potential Measure Based on Raven Progressive Matrices.

Past work with a learning potential measure using the Kohs Block designs in an enlarged format demonstrated that there was consistent differentiation of children in special classes for the educable mentally retarded (EMR) above the age of 12 or 13 years. However, below these ages there was poorer differentiation: a disproportionate number of subjects fell into the nongainer or low able category. The block design procedure also discriminated against girls and black students in a similar manner (Budoff & Corman, 1972). It was felt that the visual organizational demands of the task may be too difficult for the younger Ss and working with blocks may not motivate girls sufficiently. The literature suggests that black students, particularly Southern-born, tend to do equally well (or poorly) with verbal and performance tasks. Whether this lack of facility with performance tasks accounts for the poor showing on Kohs Block Designs of the black sample is not clear from the data. Also, the black sample was disproportionately weighted for girls. These two factors working together may have accounted for the disproportionate number of nongainers among the black sample. In part, as a function of these biases, and in part to avoid the possibility that learning potential would be determined only by one measure, re-enacting the problems connected with the IQ, we decided to develop an additional learning potential measure that would be useful for younger subjects and might evidence fewer biases than those in use with the Kohs Block designs Measure.

Initially the Research Institute for Educational Problems (RIEP) explored the utility of Feuerstein's procedure (1968) which uses the Raven's Progressive Matrices Test (Series A, A_B, and B; Form 1956) in a pre-posttest design with training on double classification problems interpolated between the two test administrations. The difficulty is that Feuerstein trains to the six 2 x 2 analogy problems which are presented in the B series and then tests responses to them on the posttest. The improvement demonstrated by trained Ss could well be due to the trained items. Also, we found the verbal instructions the principal investigator (Budoff) obtained in Jerusalem were too vague and non-specific to provide for a relatively standardized administration. The instructions did not communicate to some normal 8th grade pilot Ss what was required of them.

In lieu of being simply able to replicate Feuerstein's (1968) procedure, it was decided to re-examine the manner in which the effects of training might indicate improved ability to solve the problems presented on the Raven Matrices task. Two goals were set. The first required that the task discriminate among primary grade, lower-class children who had severe difficulties in school learning and had low scholastic aptitude scores (low IQs). The second was aimed at the teenage sample who had been examined on the Kohs learning potential procedure with the intention of developing a second learning potential procedure.

The initial experimental plan involved designing a more explicit training procedure of problems other than the double classification problems used by Feuerstein.

Hurtig (1959) and Jacobs (1971) have shown that double classification problems are too difficult for younger primary grade children even following training. To make the procedure applicable for the younger student, training procedures were added relevant to the easier items on Sets A and A_B (1956 series), namely, pattern completion and orientation so as to provide problems within the grasp of young children (on which they could improve following training).

For the older children (above twelve years) the hypothesis of this first stage of our work was that if training on the double classification problems which occupied more than half of the training procedure was effective, the Ss would show spontaneous transfer to 3 X 3 problems (Raven Sets C, D, E, 1958) for which no specific training was given.

1. Description

The Raven Progressive Matrices, Sets A, A_B, and B (1956) and Sets C, D, E (1958) are group administered before and after training. The interval between training and posttest administrations is uncontrolled because the tests are group administered but the training is offered individually in one or two sessions, depending on the child's age and responsiveness. A special answer sheet was designed which allowed the children more space for their answers. The colored matrices booklet (1956) is distributed first and completed by the children. Than a separate recording sheet for Sets C, D, E (1958) is distributed along with the problem booklet which presents 3 x 3 analogy problems in black and white designs. The tests were administered as prescribed by Raven (see Manuals). Unlimited time was allowed for each set of items, the class proceeding through the colored matrices one set at a time, waiting for most of the children to complete a set so all may start the next series together. On the C, D, and E series, the children were instructed to continue until they completed the entire book after the start of Set D. Since many of the problems in this latter booklet were much too difficult for many of the children, younger children (under 12 years) were urged to do as well as they could. They were also told that the problems were much too difficult for them and that they should not dwell too long on any one problem if they could not figure it out since they were for much older children and adults.

Throughout the training period, a number of principles were adhered to:

a. No problem which appears in Raven's Matrices is presented in the training series so that there is no specific training towards the specific test problems.

b. Great efforts are made to be certain that the child understands the demands being made of him on the task. Hence, the training procedure is not a strictly standardized one but allows for a number of variations to ascertain whether the child does understand what is required before the conclusion is drawn that he is unable to master the particular type of item. Problems will be presented in up to three different formats before training is discontinued.

c. The requirements of the task are demonstrated initially on materials that are pictorial and meaningful, ie., incomplete American flag or a Campbell soup can for the pattern completion tasks. Successive problems in a training series require the child to deal with geometric forms in order to attune the child to the basic format of the matrices test presentation.

d. The requirements of the task are concretized in motoric performances, e.g., having the child draw the item that completes the pattern before perusing the various alternative solutions presented under the problem. In the double classification problems, it was found that children could easily derive one attribute at a time but often did not hold the first attribute in mind while they derived the second relevant attribute. During development of the training procedure, the child's understanding was facilitated by having him draw the relevant attributes, one at a time, as he derived them. This helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems mentally with very little trouble.

Drawing each attribute in turn was adopted as a routine part of the training procedure. For example, when the child decided the shape of the missing figure was a square, he was asked to draw it. Then he derived the second attribute which might be dots throughout the whole figure, and was asked to add them to his drawing. The child's attention was focused wholly on the elements of the problem since the choices were covered. When the correct response was drawn, the child's attention was then directed to the choices in the lower half of the page, drawn in the same format as the Raven's test booklet, and was asked to point to the correct response for the problem. The solution process throughout this part of the training was verbalized by E.

After the child had drawn in the elements of a series of items correctly and E felt the child understood what was required on the double classification problems, the procedure was altered. In the first alteration, the child was allowed to solve the problem mentally with the alternative choices covered, and then he was asked to select the correct choice. After successful handling of this modification, the child continued solving problems in the usual mode, i.e., with the choices not covered. The child's attention continued to be directed to the elements of problems before the alternative choices were considered. Some children, in their haste, became careless and were asked to again draw the elements of the correct figure for a number of items.

It should be noted that when a child has difficulty drawing the designs, E draws them for him.

e. During the training on the double classification problems, after the child ceased drawing the alternative and began doing the problems mentally, the child was required to indicate verbally how he arrived at his solution on every third problem. This seemed to help the child maintain the quality of the solution process attained by the drawing and E's accompanying verbalization. A minimal verbal response often in combination with pointing responses by the child sufficed to indicate to E the child was solving the problems. (A more complete description is available from Budoff & Litzinger, 1973.)

2. Reliability of the Raven Learning Potential Test

Test-retest reliability for regular and special class students without training, following a one-month interval, was .817 (N = 84) and .514 (N = 45) respectively. The comparable correlations with individually administered training interpolated during a one month interval for special class students was .805 for Sets A, A_B, B ($\underline{N}$ = 403); .741 for Sets C, D, E ($\underline{N}$ = 266); and .836 for the total test ($\underline{N}$ = 259). Comparable correlations for regular class students ($\underline{N}$ = 69) were .750, .705, and .756, respectively.

What is particularly interesting is the contrast between the comparable levels of relationship between the two administrations of the test for the regular class students with or without training, whereas the special class students' scores were highly variable on successive administrations without training. Training stabilized the scores of the special class students. Test-retest reliability without training indicated very variable performances($\underline{r}$ = .54). Variability of response of these children appears to be a major problem presented by these children (Baumeister & Kellas, 1968) and provides further justification for training based models of assessment. That is, variability of response to a task will be reduced considerably and understanding enhanced, when the demands of the task are more clearly comprehended, especially in a "hands on" physically involving, competence-inducing training experience.

The data from the Raven Progressive Matrices obtained in our studies and the literature indicate the Raven Progressive Matrices has good reliability, is interesting to children, taps "g" type reasoning behaviors within a non-verbal format, can be successfully group administered as a test to primary grade and low IQ (special class) children, and correlates well with other indices of reasoning ability. These characteristics indicated that the Raven Progressive Matrices test procedure should become the primary instrument for measuring learning potential.

Following the success of the group training procedure on the Series Learning Potential Test, the training for the Raven Learning Potential task also was redesigned into a group format. Study 1 in this report reports the effects of group versus individually administered training.

C. The Series Learning Potential Test.

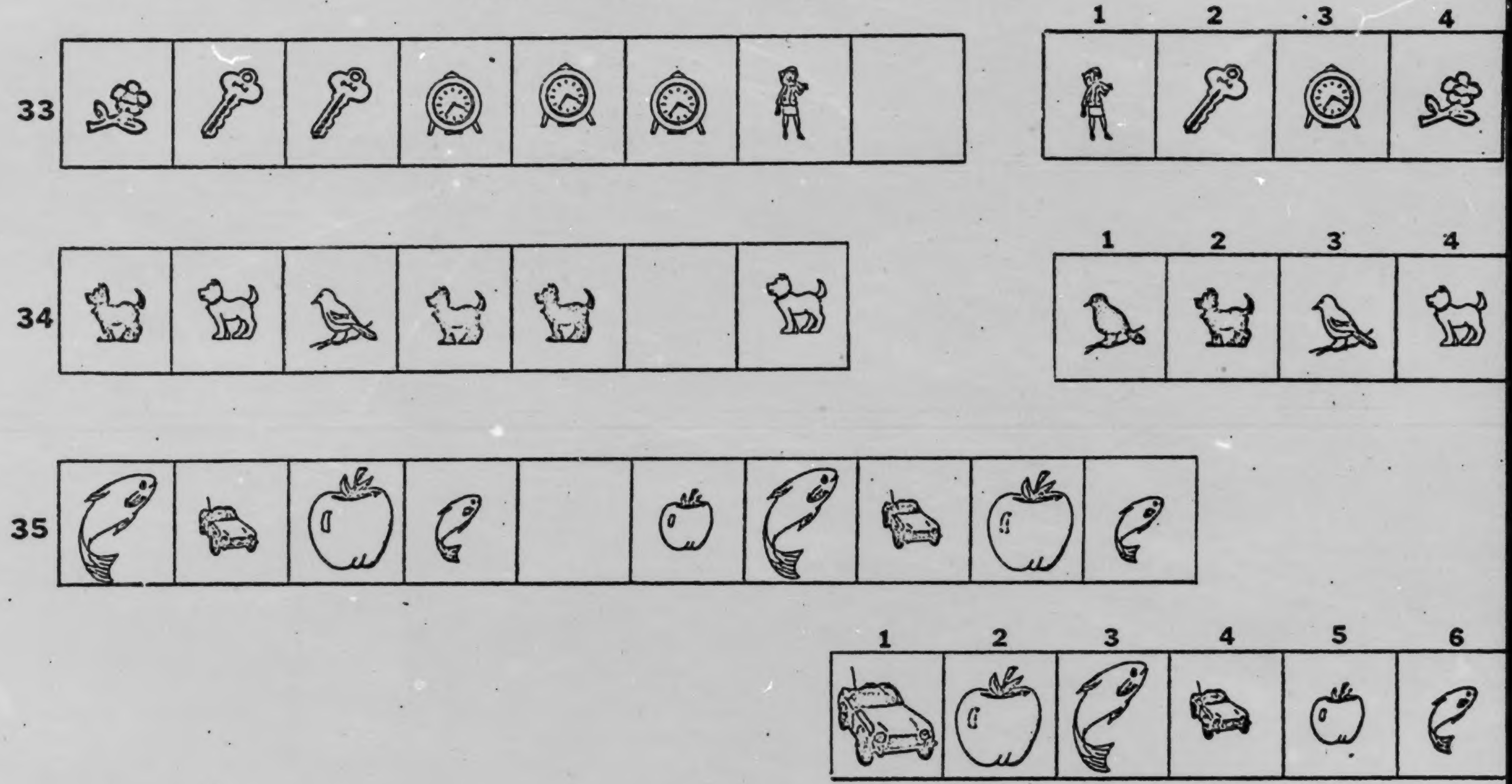
1. Description of the test.

The first entirely group-administered learning potential measure was the Series Learning Potential Test (SLPT) which was wholly developed at RIEP by Dr. Elisha Babad, now at the Hebrew University School of Education (Jerusalem, Israel). It is also a nonverbal reasoning task. Each item presents a horizontal row of cells each of which contains a stimulus figure (see Figure 3). One cell is left blank. The subjects must identify among the multiple choices on the right hand side the picture which best completes the series. Four concepts may vary in a series: semantic content (meaningful or geometric figures up to 6 instances in one item), size (large/small), color (black/white), or orientation (up/down, or left/right). Three concepts may appear in one series and they may vary symmetrically or asymmetrically. The blank space may be interposed in any part of the series. The hardest items are those in which the blank space appears early in the series, and the concepts change asymmetrically. Thus, a blank cell may require a large, black puppy dog. The child is trained to identify the separate concepts composing a series, trace the pattern of occurrence of each concept separately by "singing the tune" and cross out the choices that are inappropriate for a correct solution for that concept. Each concept is traced successively until one choice remains as the one correct answer. Thus, the child will be shown to check the size of the figure, using the "tune" -- large, large, small; large, large, small; _____, large, small. All the small figures among the answer choices would be crossed out on the answer side of the page. The child would do the same with color, "singing the tune" for the concept (black, white, black, white, black, _____) and then for the type of figure. The children are trained on items which only include meaningful pictures.

Two equivalent 65-item forms of the test are used, with items corresponding in concepts but not in specific pictures or their arrangements. The test booklets are titled, "The Series Game." None of the 17 coaching items is identical to any of the 130 test items. The test contains several types of items. Forty items consist of meaningful pictures ranging from easy items (two concepts are arranged symmetrically and the blank space appears at the end of the series) to more difficult ones (asymmetric alternation of three concepts, blank space in initial position). Improvement is solving the 40 meaningful picture series on post-test indicates the child can solve nontrained instances of the trained items. Ten series items use geometric figures to discern whether S can transfer the strategies learned in training with meaningful pictures to dissimilar stimuli.

Fifteen double-classification items are presented in matrix format to test generalization of the strategies learned to

Figure 2: Sample Series Learning Potential Items



33
1
2
3
4
34
1
2
3
4
35
1
2
3
4
5
6

problems that require the same reasoning process as the series problems, though the arrangement of the stimuli is quite different, and the strategies may be mediated differently. Five matrices are presented in meaningful pictures, and ten in geometric symbols. Table 1 summarizes the frequency of each type of items in the Series Learning Potential Test.

TABLE 1

Frequency of Items Included in the Series Learning Potential Test

Type	Frequency	
A. Pictorial series	40	
B. Geometric series	10	
C. Pictorial matrices	5	
D. Geometric matrices	10	
Total series (A & B)		50
Total matrices (C & D)		15
Total pictorial items (A & C)		45
Total geometric items (B & D)		20
Total number of items in test	65	

As developed, the test was considered suitable for children from 7 to 16 years of age. It seems now to be most appropriate for children up to the 4th grade (9-10 years) since there appears to be a de facto ceiling. In general, the pretest takes about 30 minutes, the coaching session on the next day lasts about 30 to 35 minutes. It may be varied depending on the extent to which the children require help in understanding the task. The posttest is administered on the day following training, and not later than one week following training, and takes about 25-30 minutes.

2. Description of training materials and strategy of training.

Seventeen series consisting of meaningful pictures are used in the training session. The consumable booklet is entitled, "Tricks for the Series Game." The training session is conducted in a group session. The group may range from 5 to 30 children, depending on the children, their age or ability level, and the class setting. It provides "tricks" for approaching the picture series problems. Emphasis is put on the following strategies:

a. In each series "something changes," and the child must first identify which aspects of the problem are changing (object, size, color, etc.).

b. The concept which changes in each series has a "tune," or "rhythm." After one identifies the concept, the child is shown how to discover the tune and sing it to himself. The child is encouraged to "sing the tune" beyond the blank cell in the series, and beyond the series itself. The intent is to have the child understand the regularities of the pattern.

c. Most series have more than one "tune," and the child must find all the "tunes," but "sing" them one by one. This strategy is crucial, as here the subjects learn to simplify the process of solving the problem by isolating the concepts and solving each separately.

d. For every "tune" one sings, S crosses out the choices which do not solve the series problem. One doesn't have to "remember" anything. The remaining choices show him that there are more "tunes" he must "sing."

e. "Tunes" do not necessarily start at the beginning of the series, and they can be "sung" in any direction.

In short, the children learn to form the concepts of the series using a concept of rhythm, and learn to simplify the problems by isolating the concepts, solving each concept separately. Subjects must overcome the tendency to look at more than one concept at once, or ignore irrelevant concepts. Reference to Figure 3 indicates the format of an item. Sequentially, the child is asked to identify the different ways in which the pictures change. Then, each concept, e.g., big-little, is sung by the group separately. The children are asked to leave the names of the objects to the last concept on the assumption that this would focus them on the other, less immediately obvious, attributes.

3. Construction of the test.

The equivalent pilot forms of the pretest and posttest were administered to 289 subjects. Each form was administered prior to and following training. Of the 289 subjects, 61 were classified educable mentally retarded in special classes, while the remaining 228 subjects were students in regular classes of several New England schools. Ages ranged from 7 to 16 years, although most regular class subjects were in the middle elementary grades.

All subjects were ranked according to test scores, and the test protocols of the upper and lower thirds of the sample were item-analyzed for both forms separately for the pre- and post-training administrations. The item-analysis evaluated the internal consistency and level of difficulty of each item and the amount of change in percent passed observed for each item from the pre- to the posttrained sessions.

In compiling the final forms of the test, several issues were considered. First, we wished to select items that showed good discrimination among subjects, as evidenced by the magnitude of the correlations between item and total test score. At the same time, we wished to include items at all levels of difficulty even though we knew that very easy and very difficult items may have less discriminatory power than items in the middle range. Since the posttraining performance on the Series Learning Potential Test is the crucial score for evaluating learning potential, we wished to have the more discriminatory items in each equivalent pair placed in the posttest. These items were defined as those items which showed higher internal consistency with the total test, a higher gain from pre- to posttest, and that each of the wrong choices were equally attractive.

The final forms of the test contain 65 items each. The level of difficulty of both forms is equal; 70.81% for the pretest and 70.97% for the posttest were passed by the pilot test sample. These figures are quite high due to the fact that the majority of the subjects in this analysis were bright, middle-class subjects in regular classes. The average correlation coefficient between item and total test score was .43 for the pretest and .48 for the posttest. The average change in percent passed following training for the items placed in the final pretest form was 8.72%. The average change following training for the items placed in the posttest form was 12.62%. Thus, although the two forms are equivalent in the characteristics of the items and the levels of difficulty, the posttest form contains items of greater discriminability which show greater rates of improvement following coaching.

Following this initial development work, the Series Test was used with and without training with large numbers of children in regular and special classes, ranging in age up to 16 years. The most disturbing feature of the data from these studies was the evidence for both EMRs and regular class students that the students attained a de facto ceiling on the test by 9 years of age for both the pre- and posttraining scores, though the latter scores were higher on the average and the scores were not at the actual ceiling for the test. This apparent ceiling limited the value of the instrument to primary grade children.

4. The Series Learning Potential Test (SLPT): A further test and validation (Corman).

The de facto ceiling at age 9 in these early data necessitated further study of the characteristics of the two forms of the SLPT. A younger sample was utilized to examine the curve of increase in scores as a function of CA. The broad purpose of the present study was to determine the applicability of the Series Learning Potential Test, developed at the Research Institute for Educational Problems, as a measure of learning potential.

Specific purposes of this study were to determine

a. test characteristics of the two Series forms
b. differences in initial performance by grade level (grades 1 to 4) and town
c. effects of training
d. the relationship of IQ, race, and social class to performance on the Series Test
e. the relationship of reading and vocabulary scores to scores on the Series Learning Potential Test.

Method

The Series test was given to first through fourth grade students in 79 classrooms in six Connecticut towns. All students were tested twice. Students were assigned to subsamples which determined the test forms they received and whether or not they were trained between the two test sessions. Other data collected included IQ scores, information on race and social class, and scores on the Gates-MacGinitie Reading Test. All testing took place in winter and spring of 1972.

To assess test characteristics, item analyses and factor analyses (principal components and varimax rotation) were done for each form before and after training, test-retest reliabilities were calculated, and inter-form correlation coefficients were obtained. Differences by grade and town were analyzed by two one-way analyses of variance and by graphs showing pre- to post change by grade for the middle and working class towns. To assess training effects, a four-way repeated measures design was used in which between subjects factors were grade and training group and within subjects factors were test form and test session.

Multiple regression equations with Series scores as the dependent variable were used to determine the relationship of these scores to race, social class, and IQ. These latter three variables, along with Series pretest and posttest scores, were used as independent variables in two sets of additional multiple regression equations, to predict vocabulary and comprehension scores on the Gates-MacGinitie Tests.

Results

Analyses of test characteristics showed the K-R 20 of both pretest forms to be .95 and the inter-form correlation coefficient between total pretest scores to be .84. Test-retest reliabilities were .87 and .90 for Forms A and B respectively. Factor analysis indicated a consistent pattern of items on both forms; only the Picture Series Subtest clearly emerged as a factor. The four factors obtained accounted for approximately a third of the variance in each analysis.

Analysis of variance on Series scores by town indicated that Ledyard, Wilton, Westport and Darien students had higher mean pretest scores than Stamford and Hartford students on the forms they received. These can be ascribed to differences in social class and race (former towns are white and middle class, while the latter samples were drawn from schools located in poor, largely black neighborhoods). Analysis of initial performance by grade revealed that fourth graders got the highest scores, third graders the next highest, second graders the next highest, and first graders the lowest.

Results of the repeated measures analysis of variance indicated that trained students achieved a greater mean score increase than that attained by nontrained students. Training was found to be most effective for grades 1 through 3. Graphs showed that trained second and third-graders in the communities of Hartford and Stamford (the poor, largely black samples) had the largest increase in mean scores. In the middle class towns (Darien, Ledyard, and Wilton), the greatest increase occurred for first and second graders. Fourth graders appeared to profit least from training, suggesting that the test was not suitable for children of this age or older.

Results of multiple regression analysis indicated that students who had high IQs, were white, and whose fathers had middle class occupations tended to receive higher scores on the Series tests, both before and after training. Training, however, was able to boost the mean scores of lower class students to the pretest level of middle class students in four of the five towns analyzed. (The exception was Wilton where the discrepency between initial score of lower and middle class students was the greatest.)

Correlation coefficients between Series scores and Gates-MacGinitie reading subtest scores showed that scores on these tests were positively related. The correlation coefficient between the Series pretest and posttest scores and the reading comprehension subtest were .471 and .592, respectively; with the vocabulary subtest, .326 and .429, respectively. Results of multiple regression analysis, however, indicated that social class and race were related to both subtests as well as to Series scores, so that the relationship between Series and Gates-MacGinitie scores was reduced when social class and race were taken into account.

Discussion

The Series test forms are reliable and equivalent, and training is effective for first through third grade students; however, because of the ceiling of scores at the fourth grade, and because of the relationship of Series scores to race and social class, the Series test appears to be less useful as a measure of learning potential than other measures.

The results suggest that if the scores following training were used as the index of the poor child's ability in the primary grades, a more optimistic picture would emerge of the child's capabilities. Teachers might become more encouraged by this formal evidence of the child's ability, and this might help the child's progress in school. Work in schools also suggests, however, that the school programming for children from poor homes who do well on the Learning Potential test must be reorganized for teachers to see the ability evidenced by the child on the Learning Potential test displayed in his school achievements.

Following this study, further development and use of this instrument was halted. Dr. Babad of the Hebrew University, Jerusalem, Israel, is continuing further work with this measure and has adapted it for use with older children.

Validity of Learning Potential Assessment

A synopsis of the findings relating to the validity of a training-based assessment such as the learning potential procedure was presented in the overview section. More elaborate summary presentations of these findings are presented in Budoff, 1973. In the following sections, some of the findings will be presented in some small degree of detail so that the reader will have a sense of the meaning of learning potential assessment as of the start of this project.

A. The Problem of the Criterion of Learning Potential

In its original conception, Learning Potential measurement was concerned with the extent to which the initially poor performance of IQ-defined EMRs could be improved by problem-relevant instruction on nonverbal reasoning tasks. The hypothesis was that many patients in an institution for the mentally retarded and students in special classes for the retarded could profit from such instruction and perform more effectively on the reasoning task, because they had been misclassified as mentally retarded on the basis of the IQ score. The initial thrust of this work was concerned with developing a standardized procedure which would demonstrate whether optimized procedures in testing would yield less biased estimates of ability to profit from experience, or intelligence. The initial premise was that institutionalized mentally retarded patients would score poorly even on nonverbal reasoning problems prior to training. Some proportion following training, however, would improve their performance markedly (gainers), while others would not (nongainers). Optimizing the test procedure by creating a sense of increased comfort and competence in a context of problem-relevant strategy training would provide evidence of the ability of these children in nonacademic areas of endeavor.

The initial studies with institutionalized educable mentally retarded (EMR) samples (Budoff and Friedman, 1964; Budoff, 1967) did indicate these two primary patterns of response to the training procedure. A few institutionalized persons did seem to understand how to solve these nonverbal problems prior to training and solved difficult 3 x 3 and 4 x 4 patterns. When student samples were drawn from adolescent-aged community special classes for EMRs, larger numbers of students solved the difficult nonverbal puzzles prior to training. Proceeding from easy to harder items, they were able to learn how to solve the problems much as higher IQ students performed. These students with low IQs were labeled high scorers, and most clearly represent instances of misclassification as mentally retarded.

During this early work, three patterns of response were identified on the Kohs Learning Potential task (KLPT) for those students whose IQ scores fell within the EMR range (50 - 79 IQ). High scorers demonstrated excellent understanding on the trial prior to training. Gainers performed poorly on the pretest administration, but improved their scores markedly following instruction. Nongainers performed poorly initially and did not profit from the instructional procedure. Gainers were defined on the KLPT as those students whose posttraining score was four or more designs higher than their pretraining score. This was three times the mean gain evidenced by a practiced but nontrained sample (Budoff and Friedman, 1964). Nongainers failed to attain this criterion.

When work was begun with the Raven Learning Potential Test (RLPT), the same learning potential categories were defined, but by that time the conception of learning potential had undergone some development. Reliance on gain or improvement following training posed difficulties because there was no anchoring point. To what did improvement refer? Did the child who solved no block designs prior to training but could solve four simple 2 x 2 puzzles following training possess more of the ability we characterize as learning potential than the child who solved more difficult block problems initially but did not improve his scores following training? Conceptually and empirically, we found difficulties in this formulation which was based solely on a gain score without reference to pretraining level. For the first studies with the Raven Learning Potential test, we began to work with a norm-referenced criterion. High scorers were defined as those students scoring at or above a criterion set at 2/3 S.D. below the mean score for the child's chronological age group prior to training. Gainers were defined as those who were below this level before training but who attained this level following training. Nongainers' scores following training remained below the cutoff point. The cutoff point was taken as equivalent to an approximation of a 90 IQ, which set a more stringent criterion of mental retardation than the customary cutoff of either 75 or 79 IQ. This was done to demonstrate, more dramatically than necessary perhaps, that many students in special classes for the mentally retarded were competent nonverbal problem solvers when testing conditions were optimized. One difficulty with this solution was that the cutoff point was arbitrarily established. The problems of defining an appropriate criterion of LP status for the RLPT remain poorly resolved.

More recently, we have continued to use the norm-referenced criterion, but have chosen to abandon the previous trichotomous definition of LP status. Our past work has clearly demonstrated that high scorers, gainers, and nongainers constitute a linear continuum of ability. Consequently, in our most recent analyses we have used three scores as continuous variables in regression analyses, to define three types of performance derived from the LP assessment.

Pretraining scores reflect the present functioning ability of the child on the task, and these scores correlate with verbal and performance IQs. Like IQ scores, they correlate also with such socioeconomic indicators as size of family, degree of intactness of family, race (black, white), English language competence, Binet and verbal scale (Wechsler) IQs.

The posttraining score includes three major components. First, it reflects the initial level that the subjects demonstrated on the pretest. Second, the repeated administration adds an effect due to repeated exposure to the test materials (practice). The third component of the posttraining score is the effect of training. When the systematic contribution of practice is controlled by incorporating a nontrained group in the research design, the practice effect can be estimated by noting any increments in the test-retest score of the nontrained group. Whereas low IQ children tend to gain little without specific instruction (Babad and Budoff, 1971) the third component, final level of attainment which potentially able children show, regardless of their initial score, is the most crucial. Following suitable training, many low IQ children will function at a level similar to the child from more privileged circumstances (Babad and Budoff, 1971; Corman and Budoff, 1973). This posttraining score, regardless of pretraining level, represents the child's optimal level of performance following an optimizing procedure. It permits a comparison between his presently low level of functioning, as indicated by his IQ, and his potential level of functioning. This posttraining score is hypothesized to be related to performance on tasks or curricula which permit the child to operate in areas of conceptual or cognitive strength, as opposed to areas of weakness such as the verbal conceptual domain to which IQ scores relate. It is assumed that an educational plan can use this evidence of conceptual strength to help the child upgrade his skills in the verbal conceptual areas, so that the IQ-based prognosis of failure can be reversed. The critical input is the evidence of the child's capability, i.e., his ability to profit from experience, which should set in motion teachers' expectations of more successful learning for the child.

A third score is the posttraining score adjusted for pretest level. This residualized score indicates the child's responsiveness to training, and by extension to the classroom regardless of his pretraining level. It is hypothesized to indicate the student's amenability to training given suitable curricula and school experiences. Little attention has been given this score.

Validity data for the learning potential assessment will be presented in summary form for each domain of variables in the sections that follow. The strategy for estimating the validity of learning potential as contrasted with IQ scores utilized a construct validity approach. That is, since the children of concern were poor students, academically, the usual criterion of school success utilized in validating IQ measures was precluded. The argument demonstrating the superior predictive value of a training-based score had to be constructed tapping aspects that would indicate performance in nonacademically related, real life areas of the child's current functioning requiring the use of cognitive related capabilities.

Another major strategy was to demonstrate within-group differences among children in special classes for the educable mentally retarded associated with differences in response to learning potential training, and trying to relate these differences to data relevant to low income, educationally retarded children, as opposed to the expected characteristics of mentally retarded children. The major portion of this data is derived from studies in which the Kohs Learning Potential Test was the criterion task and the learning potential status trichotomy, as derived from the gain score criterion was used. When the study used other learning potential tasks, or when the posttraining score was used as a continuous variable to indicate LP, this information will be indicated. The social-demographic and psychometric correlates of LP status will be presented for some studies in which the trichotomous LP classification was used, and for other studies in which scores derived from LP assessment were considered as continuous variables. Sections which follow will present data relevant to the educational implications of LP status, paired-associate learning, motivational and cognitive style variables. Also, capability as viewed by the students and their parents will be presented.

B. Social, Demographic, and Psychometric Correlates

Social and Test Correlates of Kohs Learning Potential Status in Community EMRs (Early Analyses)

Social history and test correlates of learning potential status based on the school records of a community EMR sample of 383 subjects drawn from a variety of communities in Massachusetts indicates some demographic and test characteristics of the three learning potential groups.

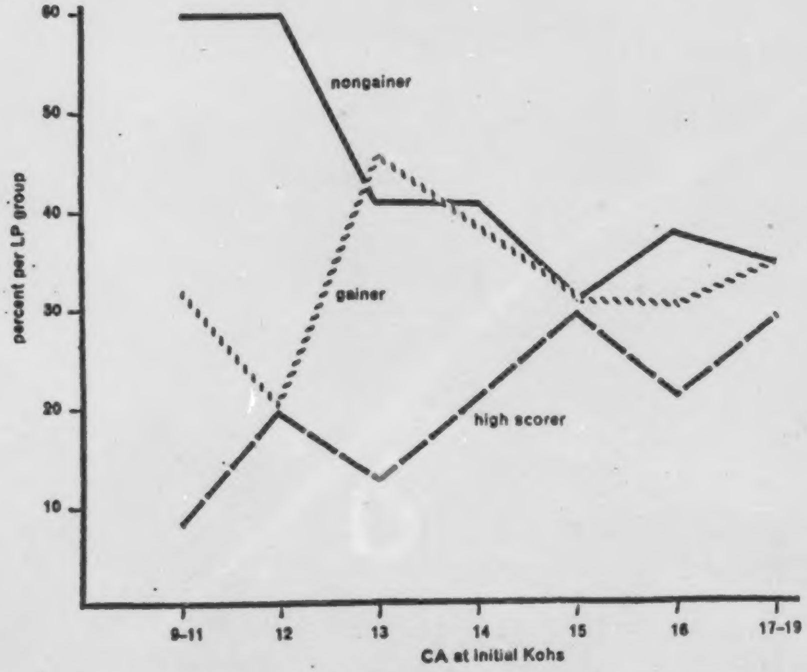
Of the total sample, 60% of the special class children born in the northeastern part of the United States attained gainer or high scorer status; 40% were nongainers. One has a maximal probability of attaining gainer and high scorer status if one is white, over 15 years, male, and from the lower socioeconomic groups and living in situations that have been classically associated with poor school performance, e.g., if one is a member of a family of recent immigrants, one's parents do not speak English, or one's family is large. The suggestion is that when the child is not from the middle class, but his family is at least benign emotionally, the child will tend to demonstrate ability on the nonverbal learning potential task. When there are instances of family instability, e.g., divorces, separation, desertions, the child has a higher probability of being a nongainer. It may be that as he grows older, he may develop sufficient ego to function more adequately on the particular learning potential task used.

The middle class child found in special classes, by contrast, tends to be a nongainer. In this instance, where there is no cultural disadvantage by which to explain the lack of school success, this inability may be due to the effects of very strong negative emotional pressures because of the parental expectations for satisfactory achievement, a "true" (organic) retardation, or an interplay of the two sets of factors.

Also, cases with clear evidence of brain insult but not including those with a history of seizures, tend to be nongainers ($X^2$3df = 8.9, $p < .03$), though other diseases - heart, kidney, tuberculosis, rheumatic fever, visual or hearing problems are not related to learning potential status.

Girls tend to be over-represented as nongainers, possibly because the girls' later placement in special class results in more experiences in failure, prior to this placement, and/or because they may respond less well to the type of problems presented in this particular learning potential assessment procedure.

Learning potential status on this task is correlated with chronological age. The incidence of nongainers from 12 years to the post school years (16 - 19 years) decreases from 60% to 33%. By contrast, about 1/3 of the samples at each age from 9 - 19 years are gainers. Because of the difficulty of the task, high scorers represent only 1/12 of their chronological group below 11 years, but about 1/3 of the post-fifteen-year-olds (see figure below).



Percentage distribution by learning potential status of special class students (n = 383).

Figure 3

These findings are in essential agreement with the correlates of learning potential status reported by Budoff (1967) with late adolescent institutionalized EMRs.

When response to the LP training task was used as a continuous variable, a step-wise multiple regression analysis was employed to determine the factors that accounted for significant portions of the variance in Kohs scores before and after training and to compare these factors with those related to scores on the Stanford-Binet IQ test.

The sample consisted of 627 EMRs from nine cities and towns in Massachusetts. Seventy-five percent (N = 471) were students in segregated special classes in public schools. Most of the remainder (N = 134) were residents of state institutions for the retarded; 22 were participants in a community workshop. The subjects ranged in age from 8 to 40 years, with a mean of 14.55 years (SD = 2.75) at the time of initial testing. Fifty-nine percent were males, 38% were white, and 79% had fathers who were manual laborers or menial service workers. Stanford-Binet IQ scores were obtained for 535 subjects; those scores ranged from 65 to 98, with a mean of 68.81 and standard deviation of 10.26.

A number of stepwise multiple regression analyses were performed against four main dependent variables: Stanford-Binet IQ, pretraining score (K_1), immediate posttraining score (K_2) corrected by pretraining score, and delayed posttraining score (K_3) corrected by the two prior Kohs scores. Major background predictors were evaluated both for their simple relationships to these four dependent variables and for their unique contributions to the variance of the dependent variable.

Sets of independent variables in the equations included chronological age, race (black and Caucasian), social class, sex, birthplace, family size (number of children) and intactness, number of diseases, and scores on various psychometric measures. Partial correlation coefficients indicated that the Stanford-Binet score was more highly related to noninstitutionalization, family intactness, and WISC Verbal IQ than were the pretraining Kohs (K_1) scores.

With all variables held constant, factors related to immediate effects of training on the Kohs were male sex; family size; birthplace outside Northeast U.S., or in foreign country; and high scores on the Stanford-Binet, Raven Progressive Matrices, and WISC Performance IQ tests. The only factors significantly related to delayed posttest, one month following training in the Kohs (K_3 with all variables partialled out), were WISC Performance IQ and scores on the Raven Progressive Matrices. WISC Verbal IQ was not found to be uniquely related to any of the three Kohs scores. In several equations where race and social class were used as independent variables, these two factors were not found to be significantly related to Kohs scores.

Scores following training did not elicit systematic differences in scores from children of different racial and social class backgrounds. These relationships were evident for pretest Kohs, Stanford-Binet, and Verbal IQ scores. When continuing effects of training are considered, the Kohs procedures are not differentially affected by age or sex as well. Performance IQ, whether measured by the WISC or the Raven Progressive Matrices, was consistently related to Kohs scores at all three points in time. The continuing positive relationship to Performance IQ reflects the nonverbal nature of the problems on which training occurred. The assumption underlying learning potential assessment that verbal skills are unrelated to general ability to reason demonstrated after a training opportunity on a nonverbal reasoning task, when the effects of all other variables are considered, was supported by the low partial correlation coefficients involving WISC Verbal IQ. This assumption was further buttressed by the increasingly attenuated relationship of Stanford-Binet IQ with the Kohs scores following training.

Similar analyses utilizing response to the Raven Learning Potential measure as the dependent measure and the same array of independent variables were conducted. The sample consisted of 403 EMR children who attended segregated special classes in Massachusetts public schools. They ranged in age from about 7 to 15, with a mean age of 11 years (SD ± 2). Sixty-two percent were male and 29% were black. Three-fourths of the subjects were from working class families. The mean Stanford-Binet IQ score was 70.95, with a standard deviation of 7.77. Raven Progressive Matrices were group-administerd by project staff before and after subjects received training. Kohs Block Designs were individually administered to 130 of the subjects.

A number of stepwise multiple regression analyses were performed against six main dependent variables: pretraining scores on Sets A, A_B, and B; Sets C, D, and E; and the total test; and three posttraining scores corrected by the corresponding pretraining scores. Major background predictors were evaluated both for their simple relationships to these six dependent variables and for their unique contributions to the variance of the dependent variables.

Variables uniquely related to a high pretest score on the total Raven test were: being male, being older, having fewer older siblings, having a father absent from the home, and having a high Stanford-Binet and WISC Performance IQ. Greater than expected improvement on the posttest was uniquely related only to being male, having a high WISC Performance IQ and long interval between test sessions. Comparison between factors uniquely related to performance before and after training indicated that effects due to age, number of older siblings, father's absence, and Stanford-Binet IQ, which were related to pretest performance, were not uniquely related to improved scores on the total Raven test after training.

Regression analyses with posttraining scores on Sets C, D, and E as the dependent variable indicated that older children and children with fewer older siblings improved the most on these more difficult items. For the total sample, social class and being white were not positively related to any pre- or posttest scores on the Raven test.

With all other variables held constant, WISC Performance IQ was found to predict significantly all pre- and posttest scores, with the exception of posttraining scores on Sets C, D, and E. By comparison, WISC Verbal IQ was not significantly related to any pre- or posttest scores, when all other variables were partialled out. The regression analysis with the Kohs subsample indicated that Kohs score was related to improvement on C, D, E, and the total test, but not to improved performance on the easier A, A_B, B items, whether scores were adjusted by pretest scores or by all variables. The Kohs sample, however, was two years older on the average, and beyond the range of the easier sets of items.

For the total sample, improved performance on the Raven Test after training was found to be uniquely related to male sex and WISC Performance IQ. Age was found to be positively related to initial performance but was not significantly related to performance on the total test after training. Scores on the total Raven test before and after training were found to be free of influence by race, social class, and verbal IQ when the total sample was considered. The predictive power of the Kohs score on the total Raven posttest was similar to the predictive power of the Raven score on Kohs scores after block design training, as indicated by the previous analyses.

C. Educational Implications of Learning Potential Status

1. Learning on a Minimally Verbal Curriculum

Budoff, Meskin, and Harrison, (1971) developed a model for testing the educational significance of Learning Potential scores. They adapted an educational unit, developed by the Elementary Science Study (1966), that taught simple electrical concepts by allowing the student to manipulate flashlight batteries, bulbs, and copper wires. The investigators formulated a minimally verbal test by which to evaluate how much the students knew about electrical circuits prior to the course, and then retested them following teaching the unit. On this test, the students could demonstrate their understanding by pointing to electrical circuits mounted on pegboards to indicate which bulb would light in one section. In a second section, they merely had to indicate whether the bulb specified in the circuit on the pegboard would light or not, and if so, whether it would be brighter, dimmer, or of the same brilliance as the standard unit of electricity (a lit one-bulb, one-battery circuit they held in their hands). This format was also repeated in a written version when they answered simple questions on printed diagrams that closely resembled what actual circuits looked like, and, in another section, utilized schematic or symbolic representations of these circuits as they might appear in a physics text.

The evaluation of the laboratory science classes supported the argument that high Learning Potential status students are educationally rather than mentally retarded. That is, while the level of scores on the evaluation instrument did differentiate between the special and regular class students prior to teaching, following teaching, special and regular class children could only be differentiated on the basis of their Learning Potential status, not their class placement. This was most clearly evident when the performance of the low achieving CA controls was compared with the special class samples. The superior level of performance on the pretest of the low achieving CA controls was no longer evident on the posttest. There were no differences between the two groups when posttest or improvement scores were compared. Almost all the significant F-ratios were lodged in the main effect for Learning Potential. Even more consistently than with the larger sample of CA controls which included some high and average achieving students, the low achieving students' level of knowledge on the electricity test was indistinguishable from that of the special class students taken as a group, following exposure to a suitably designed curriculum unit.

The serious defects in the special class students' abilities became evident when the verbal reasons they offered for their empirical choices to 16 pegboard problems were analyzed. The understandings displayed in these two response modes were compared for the posttest. The verbalizations offered for correct answers were coded as to whether they were descriptions of the set-ups, "there were two bulbs," or whether they offered some causal explanation for the choice "the power is shared by two bulbs." These statements were qualitatively weighted for adequacy of response.

The LP status predicted the level of verbal scores across special and regular class samples when adequacy of response was considered without regard to whether the explanations were coded as descriptive or causal ($F = 5.83$, 2/42 *df*, $p < .05$). The special class high scorers performed as well as any regular class subgroup, except the special class gainers and nongainers scored well below all the regular class groups.

The critical difference between the abilities of the regular and special class students emerged when adequacy and type of response were considered together. There were no differences between special and regular class subjects in giving descriptive reasons, but marked differences in giving causal explanations for their pegboard choices ($F = 4.66$, 1/42 *df*, $p < .05$). The special class high scorers performed as well as any regular class subgroup, except the special class gainers and nongainers scored well below all the regular class groups.

The critical difference between the abilities of the regular and special class students emerged when adequacy and type of response were considered together. There were no differences between special and regular class subjects in giving descriptive reasons, but marked differences in giving causal explanations for their pegboard choices. ($F = 4.66$, 1/42*df*, $p < .05$). Thus, even though special class high scorers were correct on the pegboard items as often as any regular class LP group, they gave a good causal reason for only 29 percent of their responses, and *every* regular class LP group exceeded this percentage.

In conclusion then, what should be evident from these data was that special class children are capable of learning empirically as well as regular class children. They demonstrated marked difficulties when they were required to explain the reasons for their responses. On this latter measure, all the special class LP groups performed poorly, especially in offering causal explanations.

The contrast in the findings for the empirical or nonverbal and verbal modes of response would seem to reconcile the differences evident between a judgment of competence based on the learning potential assessment and more traditional evaluation procedures. The special class high able LP subjects' learning ability was most adequate when their competence was judged on the items to which they could demonstrate their empirical understanding with a minimal verbal response. On the traditional ability measures, they showed their inadequacy because they were required to respond with adequate verbal reasons to items on the individual IQ or achievement tests. The non-traditional LP approach showed clear differentiations within populations of children who are said to be limited in their ability. The measure, based on their verbal explanations, tends to confirm that these children's difficulties lie in explaining problems that they understood empirically. Their performance verbally does agree with the IQ-defined deficits. But the ability the more able learning-potential students demonstrated when the verbal demand was minimized was similar to those of educationally retarded children of high IQs.

D. Learning Potential and Paired Associate Learning

A number of studies within the Institute have compared the learning performance of learning potential groups on meaningful picture paired associate tasks. With a relatively easy task composed of only six pairs of dissimilar pictures, the scores of high scorers and gainers were combined and compared to nongainers. No differences were found between learning potential groups in three samples: older (CA 17 - 23) and younger (CA 10 - 13) institutionalized samples and a community special class sample (CA 9 - 17) (Budoff, 1965). With a minimal information load, Kohs LP groups are not differentiated on the usual paired associate learning task.

When the information load was increased to ten dissimilar picture pairs and the samples were composed of community special class students ranging in age roughly from 7 to 15 years, the findings are more variable. Gainers learned faster than nongainers with whom they were matched on CA, IQ, and pretraining Kohs blocks performance (Budoff, 1965). Similarly, White (1968) found gainers learned faster than nongainers, with gainers performing comparably to high scorers, when LP status was defined by the Raven Matrices Procedure. No differences, however, were found between Kohs LP defined gainer and nongainers by Glaser (1966) or Frank and Riley (1969), when the samples were not matched by starting level. The latter study did find high scorers to be superior to the other groups when LP status was defined by the Raven Matrices procedure.

When the information load of the picture paired associate task was increased to 18 items, clear differences appeared among LP groups ($\underline{F}_{2/21} = 3.58$, $\underline{p} < .05$). Analyses of the differences in number of errors, indicated the high scorers learned with significantly fewer errors than the gainers and nongainers ($\underline{F}_{1/21} = 6.93$, $\underline{p} < .025$). There were no differences between the latter two groups (F < 1). Although the mean differences were significant, there was considerable overlap of performance among groups with one nongainer performing better than all other Ss. It appears likely that with minimal information load, LP (Kohs) groups do not differ in ability to learn a picture paired associate task. When the load is increased moderately (to 10 items) there may be a point at which nongainers are outperformed by gainers who perform comparably to high scorers. But beyond this point, under conditions of overload (18 items) the performance of the gainers breaks down such that they no longer differ from nongainers, and both groups learn more slowly than high scorers.

Mankinen (1972) trained samples differing in LP status to embed each of 23 picture pairs in a sentence. One week following training, they were retested on an alternate series of picture pairs. High scorers whose performance had improved markedly immediately following training also learned the new series with few errors. Gainers made fewer errors immediately after training but learned less efficiently one week following training, a pattern similar to that displayed by the nongainers. Nongainers, however, made more errors in both posttraining trials, though they did show marked reduction in errors.

E. Learning Potential in Relation to Personality Variables

Differences in ability to profit from experience would be hypothesized to relate to personality variables. More able learning potential individuals should be more able to manage and control one's behavior to cope more adequately with stress, to be better adjusted and to be understanding and planning more realistically for one's future. The mentally retarded child has been described as unrealistic in his striving, and relatively unable to cope with stress. If, as hypothesized, the more able learning potential student is educationally retarded, not mentally, he should cope more adequately with the stresses of his environment than the less able (LP) nongainers.

The following series of studies examined this hypothesis, largely comparing performances within the special class samples as a function of LP status. An interview study compared the statements of special and regular class students in four areas - attitudes toward school, toward work and future vocational planning, social activities, and their perception of their families. The findings of the separate sections will be presented in summary form.

1. Learning Potential and Self Concept

An adaptation of the Laurelton Self Concept Scale (Guthrie, et al., 1961) and the Bialer-Cromwell Locus of Control Scales (Bialer, 1961) was administered to 172 psychometrically defined EMRs and these were factor analyzed (Harrison and Budoff, 1972a, b). Most apparently, maladjustment is correlated with a poor Kohs performance throughout. Various components of maladjustment, however, show different patterns of correlation with successive Kohs administrations. The neurotic anxiety component correlates with poor Kohs performance on the first two tests, and then drops from significance. The acting out component correlates with only a poor first performance on the Kohs. Friendlessness, on the other hand, is increasingly correlated with a low Kohs performance on successive tests, and suggests that inability to relate intimately interferes with the acquisition of Kohs skills. Alternately, this finding may indicate that those of low, less modifiable intellectual capacity have a great deal of difficulty making friends.

Other correlates of a poor Kohs performance and/or poor gain scores on successive administrations are bravado, mischievousness, obedience, and depressed self-criticism. Correlates of a generally good Kohs performance are feelings that others think well of one's physical capabilities, the fantasy that one will succeed in adversity, and high motivation for school.

The most intellectually disabled children (by the Kohs learning potential criterion) in this population also report themselves to be the most emotionally and socially handicapped. Children who project the fantasy of self-worth and those who feel that they can succeed interpersonally under difficult odds do better and improve more on the Kohs LP tasks than those who do not have fantasies of worth and success. A pattern of peer rejection and obedience at home seem to be major correlates of poor Kohs LP performance.

2. Learning Potential and Level of Aspiration

When aspiration level of gainers and high scorers was compared in an experimental situation in which actual performance was controlled, and success and failure conditions manipulated, gainers' aspirations tended to track their actual performance. Nongainers' aspirations, however, tended to be distorted under a condition of failure, and most especially, when they had already experienced success.

The three main findings of this study appear to be:

A. Ss who can be taught to do reasonably well on the Kohs blocks after training in principles appropriate to solution (gainers) are more realistic in their goal-setting behavior than those subjects whose performance on the Kohs blocks does not improve after teaching (nongainers);

B. The disruptive effect of failure on realistic goal setting is greater for nongainers than gainers; and

C. Temporal patterns of success and failure have a much greater impact on the goal-setting behavior of nongainers than on that of gainers (Harrison, Singer, Budoff, and Folman, 1972).

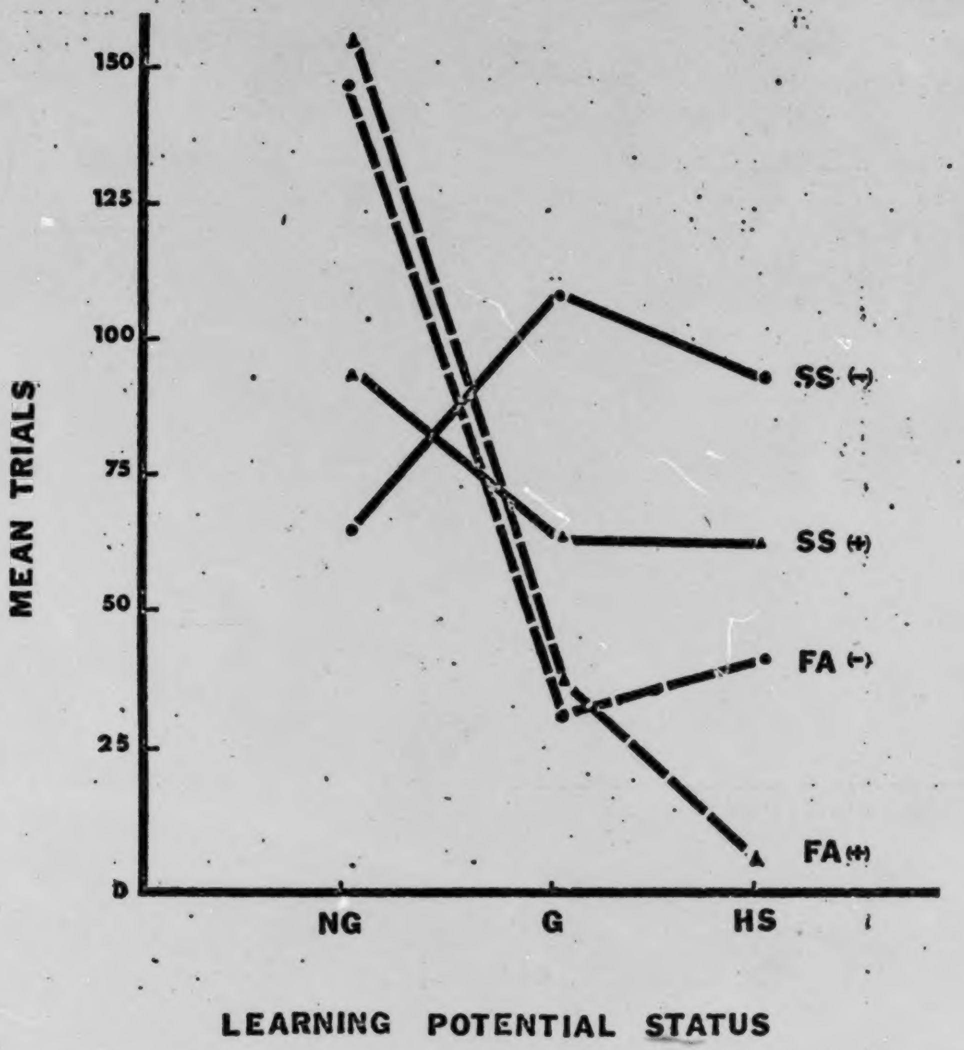
3. Learning Potential and Success Striving versus Failure Avoiding (R. Mankinen)

Forty-eight special class EMRs from an urban and a suburban school who had been administered the Kohs Learning Potential Test were classified as having failure-avoidant (FA) or nonfailure-avoidant (NFA) behavioral styles according to the relative difficulty of a block design they were willing to construct following successful completion of an initial subject-chosen design. This designation was made by asking the student in an individual session which of the block designs, with which he was already familiar, he would choose to try. Those whose choices tended to be easier than the designs they had successfully completed during the posttraining learning potential session were considered as failure avoiders, i.e., they were concerned with minimizing the prospects of failure. Those who chose designs that were at or above their highest correct solution were categorized as success strivers. The prospect of failure, per se, did not seem to govern their choice behavior on this task.

Discrimination learning efficiency was determined with a two choice simultaneous discrimination problem under one of two concretized feedback conditions: positive feedback (S received tokens for correct responses) or negative feedback (S forfeited tokens for incorrect responses). Design: 2 x 3 x 2 x 2 factorial with factors of school (urban, suburban), learning potential (nongainer, gainer, high scorer), behavioral style (failure-avoider, success-striver), and feedback (positive, negative).

Analysis of trials to criterion yielded a behavior style by learning potential interaction, $F_{(2,24)} = 8.02$, $p < .002$). Figure 4 indicates the results. Simple effects analysis revealed no discrimination learning differences among LP groups who were success-strivers. The performance of failure-avoidant nongainers, however, was strikingly inferior to that

Figure 4


MEAN TRIALS
150
125
100
75
50
25
0
NG
G
HS
LEARNING POTENTIAL STATUS
SS (−)
SS (+)
FA (−)
FA (+)

of failure-avoidant gainers and high scorers. The results indicated that nongainers are not of uniformly poor learning ability relative to gainers and high scorers and that the performance of all three groups is partly determined by behavioral styles which may be somewhat independent of actual learning ability.

4. Learning Potential and Response to Frustration

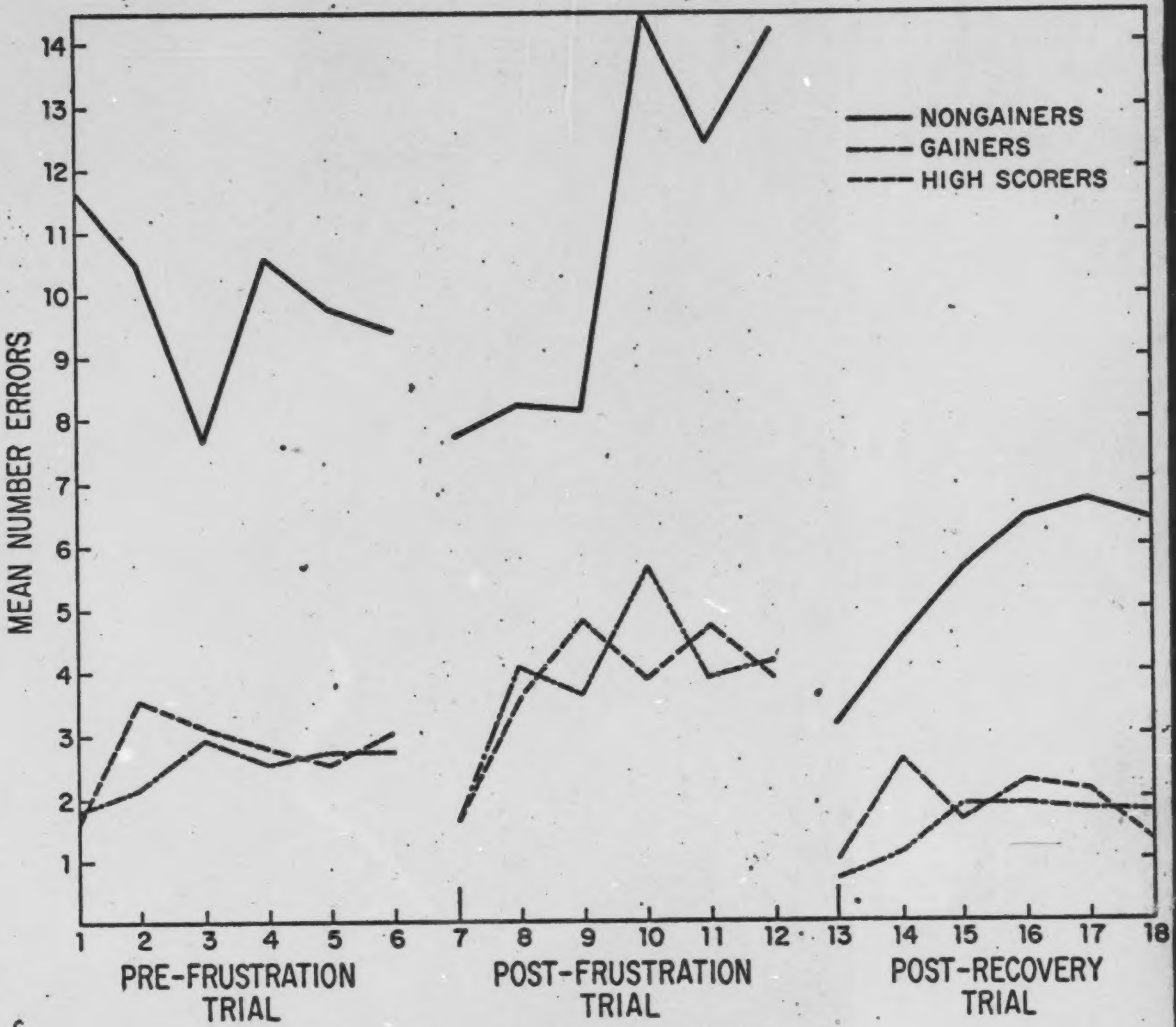
On a *behavioral* measure of response to frustration, Budoff and Pines (1970) also reported a tendency toward less disruption following a mild stressor (mirror tracing of a six pointed star) among the more able (LP) special class students than among less able nongainer peers. The criterion task was the Pins and Collars subtest of the Crawford Small Parts Dexterity Test. The students were required to place a small steel pin in a narrow hole, and a collar on top of the plate using a pair of tweezers. Facility on the task is easily subject to disruption by tension.

Learning potential status mainly predicted rate of learning on the motor task ($F = 6.814$, $p < .05$). There was support for the hypothesis that high able (LP) students (high scorers and gainers) would cope with stress more adequately than the low able LP nongainer students ($F = 2.97$, $p < .05$). As Figure 5 indicates, the performance of high scorers and gainers are essentially very similar during all three phases of the experiment. The performance of the nongainer group, however, differs in some interesting aspects. They performed erratically during the pre-frustration phase. They showed a markedly delayed response to frustration, however, unlike the other groups, and unlike the response that would be expected. They demonstrated very marked proficiency early in the post-recovery phase, but were unable to maintain this, and their performance slowly deteriorated. The pattern would suggest that these *Ss* are unable to maintain their initial appearance of adequacy (Budoff and Pines, 1970).

A *questionnaire* and *sentence completion test* which consisted of hypothetical frustrating situations in the areas of school, peer and parental relations was administered to educable mentally retarded (EMR) groups of nongainers, gainers, and high scorers, defined by the Kohs learning potential (LP) procedure. All the responses were designed around the central theme of the children's response to hypothetical frustrating situations. The responses to each of these measures were analyzed along two dimensions: 1) active or passive and 2) positive or negative. The activity described in the response was defined by its relation to the frustrator. If an act was directed toward the frustrator, it was coded as "active." If an act was directed toward oneself or away from the frustrator, it was considered "passive." The positive-negative dimension refers to the social acceptability of the act or the emotion involved. The two dimensions were combined into the following scoring categories:

1. *Active-Positive*: Acting in a socially acceptable way towards the frustrator.

Figure 5


MEAN NUMBER ERRORS
1
2
3
4
5
6
7
8
9
10
11
12
13
14
NONGAINERS
GAINERS
HIGH SCORERS
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
PRE-FRUSTRATION TRIAL
POST-FRUSTRATION TRIAL
POST-RECOVERY TRIAL

2. *Active-Negative*: Acting in a sociallly unacceptable way towards the frustrator.

3. *Passive-Positive*: Withdrawal from the frustrator through acting in a socially acceptable way.

4. *Passive-Negative*: Withdrawal from the frustrator in a socially unacceptable way.

Sample: The sample of 27 black special class adolescents, subdivided by learning potential status (N = 9 per group), was used for each of the measures, which were administered individually during several sessions in a long series of contacts by the same E (Pines). All Ss were residents of a low income housing project. They ranged in age from 10 years, 7 months to 16 years, 7 months; and in IQ from 69 to 80.

Results: The answers to the four categories of responses to the questionnaire were subjected to separate analyses of variance in which learning potential status (three levels) and sex (two levels) constituted the between groups variables, and content area (school, friends, parents) and form of question (open versus closed-ended), the within effects. CA and IQ were included as covariates. There were no significant F-ratios for the triple or quadruple interactions.

Learning Potential (LP). The two degrees of freedom associated with LP were analyzed separately for the linear and the quadratic components. The learning potential hypothesis predicts that the linear continuum of high scorers, gainers and nongainers, describes an ability dimension, from most to least among special class students. The linear component of the LP main effect tests this hypothesis (H_o = high scorers = nongainers). The quadratic component tests a motivational component, more to less accessible to training [H_o = gainers = (high scorers + nongainers)/2].

Summing across content categories, special class gainers and high scorers gave many fewer responses coded in the passive categories than the nongainers and many more coded "active-positive." Gainers gave the fewest passive and the most active-positive responses. (LP main effect - $F_{1/21}$ = 17.04, 10.16, 37.38, p < .01. for the linear component, and 8.50, 15.85, and 54.03, p < .01 for the quadratic components for these response categories, respectively.) The quadratic comparisons are due to the effects of the nongainers, since there were no differences between high scorers and gainers. There were few active-negative responses which were given by any group, and no systematic differences by LP status.

High scorers and gainers gave more passive-negative responses to peer and family rather than school situations, unlike the nongainers who were most negative toward school. Gainers and high scorers responded more positively to school situations than the nongainers (LP_{1in} x content area interaction, $F_{2/107} = 8.59$, $p < .01$, for passive negative responses and $F_{2/107} = 3.95$, $p < .05$ for active-positive responses.)

Most active-positive answers were given to frustrating school situations, the second greatest number of frustrations at home, and the least to friends; while more active-negative answers were given to questions regarding friends, than to parents or school (least). Most passive-positive answers were given to items relating to parents, fewer to friends and the least for school items. These results make sense since clearly peer relations permit the greatest latitude, and school the least latitude for negative behaviors. There were sex differences but these are not discussed here (see Pines and Budoff, 1970).

Response to Open Versus Closed-Ended Questions. A subsidiary question concerned the consistency of responses across the open and closed-ended conditions. It has been observed that nongainers tend to be inconsistent verbal responders in an interview situation, especially as the questions became more specific. The situations were presented with responses required to systematically test this observation to the two types of questions.

There was a significant main effect for open and closed-ended responses only in the passive-positive category, probably because of the popularity of "don't care" responses among the forced choices presented in the closed-ended format (open versus closed-ended main effect, passive-positive category $F_{1/107} = 5.54$, $p < .05$). In comparison with high scorers and nongainers, the gainers gave more passive-negative answers to the closed-ended choices ($LP_{quadratic}$ x open versus closed-ended, $F_{1/107} = 4.34$, $p < .05$).

When the discrepancies in the coded choices given to the open and closed-ended questions were summed over each *S*'s responses to the 15 questions, high scorers and gainers gave the fewest discrepant choices (mean scores were 5.44 and 4.11), while nongainers' responses tended to be more discrepant (mean of 9.44; $F_{2/24} = 24.15$, $p < .001$). This finding of discrepant responding across the two modes of response to the same situations indicates the nongainer does tend to be an inconsistent responder. By contrast, the gainer and high scorer tends to stick to his answer even if, as in the forced choice situation, he realized that there was another answer that he preferred. The nongainer seemingly either did not care if he did change his answer, did not even realize that he was giving another answer, or was more flippant in his verbal responses.

Five frustration sentence completion items were presented in a format covering four content areas -- school, peers, mother and father, and were coded for the same categories. The same analyses were performed.

Again, the major hypothesis was supported. Membership in a LP status group was associated with differing types of responses. As with the questionnaire, nongainers' completions were more often passive-negative when compared to high scorers (LP_{lin} main effect, $F_{1/21} = 30.23$, $p < .001$). High scorers gave more active-negative and positive responses than nongainers (LP_{lin} component, $F_{1/21} = 5.69$, $p < .05$ and 16.30, $p < .01$ for the two categories, respectively). When gainers' responses were compared with those of the high scorers and nongainers, the significant quadratic components indicated that gainers gave fewer passive-negative and more active-positive responses than the high scorers and nongainers (LP main effect, quadratic component, $F_{1/21} = 34.94$, $p < .001$, and 25.42, $p < .001$, for the two categories, respectively). Again these effects seem mainly due to the deviant pattern evidenced by the nongainer. There was relatively little difference between high scorers' and gainers' means.

5. Learning Potential and the Stroop Color Word Test

The Stroop Color Word Test was administered to determine whether special class adolescents differing in Kohs LP status would also differ in coping with the mild interference task dissonance of having to report the colors when the color names are printed in an ink of a conflicting color. The task was administered on three successive days because Jensen (1965) has reported marked reductions in time required to complete the Stroop tasks on successive administrations of the tasks (over ten successive days to college students). Examination of his data indicates the most marked reductions occurred during the first three repetitions, after which an asymptomatic level of performance was attained. The Ss in this study were administered the tasks on three successive days to determine the effects of practice, and to allow the Ss to demonstrate a more optimal performance level. The usual instructions were employed at each administration.

Results: Learning Potential Status. The prediction that high scorers' and gainers' performances would show less interference than nongainers' on the color-word color task tended in the predicted direction and represented the only significant effect for learning potential status main effect for either the linear or quadratic components ($F = 3.07$, $.10 < p > .05$).

High scorers were less susceptible to interference effects than the nongainers. There were no significant differences in performance on the color or word naming cards, which indicates that the three learning potential groups were not different in performance on these simple tasks, but rather in a personality characteristic, namely, the ability to inhibit response to a conflictful stimulus and to respond adequately with a minimum time delay.

6. LP and Rigidity in the Adolescent EMR (Budoff and Pagell, 1968)

Lewin (1936) argued that retardates, as compared to normal children of the same degree of differentiation, that is, MA, were characterized by a lessened permeability in the boundaries between cognitive regions. Kounin (1941) extended this position by stating that the degree of rigidity in these boundaries was a positive monotonic function of CA. This argument generated the prediction that, with MA held constant, older retardates suffered from greater cognitive rigidity than younger retardates, who were more rigid than normal children. This prediction was supported by a series of five experiments (Kounin, 1941). When Ss were told by E to draw cats until satiation, and, following this, bugs, the normal children drew for the shortest period of time on the second drawing task as compared with the first, the younger retardates for a longer time, and the older EMRs for the longest period. When Ss had to initiate a change in sorting principle on a concept-shift task, the normals evidenced the least difficulty in shifting, the older retardates the most difficulty, with the younger EMRs again falling between these two groups.

Zigler (1962) has argued that the differences in performance that Kounin demonstrated may be largely accounted for by motivational factors. He cited experimental evidence from his own work which indicates thet EMRs tend to approach novel situations with a negative reaction tendency; that they perform more satisfactorily when offered tangible rather than intangible incentives, such as praise; and that they tend to be more conforming to adult demands in an experimental situation because they have experienced social deprivation in their past lives, or in their present institutional environments.

The learning potential procedure elicits different levels of performance on a nonverbal task within homogeneous IQ populations of EMRs when several of the conditions cited as critical by Zigler are included as a means of maximizing S's understanding and participation. In the present study, measures of rigidity were administered to gainers, nongainers, and MA and CA controls (N = 20/group). Gainers, because of their performance on nonverbal tasks following training were hypothesized to be educationally rather than mentally retarded. Hence, it was predicted that gainers should perform similarly to CA controls. Nongainers, who tend to perform on nonverbal tasks at levels consistent with their retarded IQ scores, were predicted to perform similarly to MA controls. Gainers were predicted to perform less rigidly than nongainer EMRs.

Zigler (1961) has shown that social deprivation in the life history of normal and EMR Ss is associated with increased rigidity. This independent variable was included in the design of the study. Socially deprived Ss were hypothesized to perform more rigidly than nondeprived Ss.

Sample: Twenty special class gainers and 20 nongainers between the ages of 12 and 17 were selected. Mean CA was 14.6; mean Binet IQ of both groups was 72.8. The MA control group consisted of a sample of 20 fourth grade children of average IQ. Twenty ninth graders of average IQ served as CA controls. All four groups of Ss were selected from working class backgrounds as determined by father's occupation and parents' educational level. Each of the groups contained 10 boys and 10 girls. The groups were further subdivided into socially deprived and nondeprived categories after consultation with teachers and school records.

Measures: Two measures of rigidity were employed: a simple "marble-in-the-hole" game, Zigler's adaptation of the problem used by Kounin (1941), and an adaptation and extension of Kounin's concept-switching task in which 27 cards could be sorted for three concepts: shape-circle, star and square; color --red, green and yellow; and number of figures -- one, two and three.

Procedure: Two sessions were held with each S individually. At the first session, simplified versions of the two experimental tasks were presented in a supportive atmosphere that would minimize apprehensiveness. Within one week, E presented the two experimental tasks. When the concept-switching task was presented, S was told he would get a candy bar when he was able to sort for each concept successfully. Seven trials were allowed for each concept sort.

Results: The time spent on the marble task and the trials required to switch concepts were analyzed separately. The mean times for the four subject groups on the marble task were significantly different. The analysis of the differences between these means indicated that the CA normals spent markedly less total time on the task than the MA normals ($p < .01$), nongainers ($p < .01$), and gainers ($p < .05$). The gainers tended to spend less time than nongainers and MA normals ($.10 < p > .05$).

There were no differences between nongainers and MA normals in mean time.

The CA normals also required fewest trials to switch concepts. The gainers required fewer trials than the nongainers ($\underline{p} < .05$), but the MA normals did not differ from either the gainers or nongainers, the mean score falling between these two groups. Ranking the means, the CA normals grouped for the three concepts most rapidly (M = 1.05 trials per concept), followed by the gainers (1.90), MA normals (2.47) and nongainers (3.30). All three groups required significantly more trials to switch concepts than did the CA normals. (See Figure 6 below).

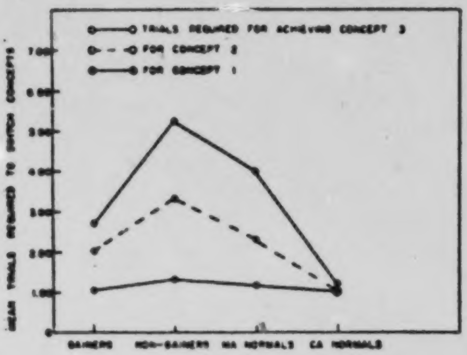


FIG. 6 Groups × Concepts interaction for concept-switching task.

An analysis was performed to determine the order of the difficulty for attaining the three concepts: color, shape and number. The mean number of trials required was 1.83, 1.86, and 2.58, respectively. The number concept was significantly more difficult to attain than color or shape ($\underline{t} = 2.0, 2.1$, $\underline{p} < .05$). When the number of children in each group who attained each concept sort was tabled, fewer nongainers sorted for shape than members of the three other groups, and fewer nongainers and MA normals sorted for the number concept (for shape $\underline{X}^2_{3df} = 9.430$; $\underline{p} < .05$); for number: $\underline{X}^2_{3df} = 12.180$, $\underline{p} < .01$; for color: $\underline{X}^2_{3df} = 6.0$, nonsignificant.

Students from socially deprived backgrounds spent more time on the marble task, supporting Zigler's social reinforcement hypothesis. However there were no differences in performance on the concept switching task which could be ascribed to level of social deprivation.

7. Learning Potential and Impulsivity

Mankinen (1970b), using an adaptation of Kagan's (1965) Matching Familiar Figures (MFF) task as a measure of impulsivity with white and black adolescent special class students reported that learning potential differentiated between the number of figures correctly identified, though there was no difference among the group means in latency of response. The more able (LP) Ss (high scorers and gainers) demonstrated less impulsivity, matching more figures correctly than the less able nongainers. ($F = 11.02$, $p < .005$).

F. Summary of Interview With Special and Regular Class Adolescents. Topics Covered Were Peer-Social, Family, Vocational, and Academic Areas.

The basis of the LP argument is that if the able low IQ child from poor and/or culturally different background is offered problem-relevant training in a comfortable competence inducing context, he will demonstrate increased proficiency on the tasks. Since the tasks we have used are nonverbal, the relationship of improved performance to traditional school-related criteria of ability tends to be low. Evidence to support the argument must be developed in curricula or teaching situations specially designed to tap these students' cognitive strengths. With a few exceptions, we have been unable to demonstrate marked changes in traditional academic achievements that depend on verbal-conceptual or expressive or reading skills even under optimized learning conditions. These types of attainments may not be possible for children who have been educated for any period of years in special classes for the retarded after two or more years of failure in the regular grades.

Evidence of this ability to learn and profit from experience that is not dependent on school types of achievements and skills should be available from the child's attainments in the nonschool-related areas of his life (Mercer, 1972). This should be especially true with poor children and their parents who tend not to devalue children who have been labeled as "dumbies" because a high proportion of them have experienced school failure themselves, or have witnessed it occur to close friends and family members. One source of construct validity for the evidence of ability derived from LP measurement is the nature and character of the life experiences of the child outside of school. The most feasible source of this information is the child himself.

An interview schedule was developed that sought information in four areas -- vocationally relevant activities - past, present and future, and social, family and academic relationships. Information was also sought relevant to their attitudes and activities in school. The general hypothesis in each area of the interview was that the high able (LP) special class students will display behavior, interests and aspirations more characteristic of the regular class respondent, while the low able LP students' responses will differ, and reflect responses suggesting retarded or deviant behavior in the particular area. Separate complete reports presenting the results in these four areas are available.

Sample: The students were drawn from a junior high school whose students are low income Irish Americans. All the students in the three special classes for the educable mentally retarded (EMR) were included in the subject pool; the only exclusions were those with demonstrable cerebral involvements, or extremely emotional disturbance, i.e., having a suspicion of psychosis, not merely a severe behavior problem. The regular class children, who served as controls, were drawn from three seventh grade sections: a low academic track section, and two low general track sections. All the children in these sections had failed two or more major subjects following the close of the first marking period. While some *Ss* would improve their performance during the school year, it was felt that these students did represent an "at risk" population.

The occupations of the mothers and fathers of the children were obtained from the school records. There were no significant differences between any of the groups by social class when this was determined by rating the principal wage earner's occupation according to Turner's classification of occupations (Turner, 1964). The occupational ratings fell into the categories Turner designates as unskilled (1), semiskilled (2), or skilled (3) laborers.

Instrument: The interview consisted of four sections with a total of 237 questions. There were two distinct types of questions employed -- an open-ended question in which *E* asked directly for *S*'s interests and desires and a multiple choice question in which *E* presented *S* with a hypothetical situation followed by several alternative outcomes of which *S* had to pick one most representative of himself. The multiple choice design was employed for the role playing, sociometric, locus of control and responsibility questions, all of which were modifications of questions used by other investigators.

On the sociometric questions, *S* was presented with a situation (e.g., social -- want to go to a movie; family -- unhappy and need someone with whom to talk; and academic -- need help with homework) and was asked for the persons he would and would not choose for that specific situation. He was required to give both a child and an adult. The object of these questions was to see with whom S identifies, whether or not his peers were from special or regular classes, and whether they changed as the situation changed.

S was given a situation into which he was supposed to project himself for the role-playing questions (club activities, household duties, class participation). He was given a choice of four roles that were available (ranging from observer to leader) and was asked to name the one for which others would choose him and the one he would like to have. Each of the roles was assigned a weight prior to the administration of the interview. The scores were the same in each section with observer = 1, passive participant = 2, active participant = 3, and leader = 4. We were trying to see whether or not S's role outside of school was the same or different from his classroom role.

On both the locus of control and responsibility questions, S was presented with two alternative situations, one responsible, the other irresponsible, and asked to choose the one most applicable to himself.

The content of the direct questions was based on areas that previous investigators had found to differentiate among school children and on areas that might be hypothesized to differentiate among special and regular class students. Each of the direct questions was preceded by a short statement, e.g., "Other things that some kids do after school or on weekends is take lessons. They may be lessons in religion, music, art, etc." The purpose of these statements were twofold. The first was to imply to S by saying "some kids," that not everyone does engage in this activity and it's perfectly alright not to do so, thereby reducing the probability of getting socially desirable answers. The second purpose of the introductory statement was to tell the subject, by giving examples of music, art, etc., to what exactly the question was referring. A more detailed description of the questions involved in each area will be presented in the individual reports on the individual sections.

Procedure: Each subject was individually administered the interview. The interview was divided into four separate sections and was presented in the following order: social, family, vocational and academic sections. The subjects were not required to read or write anything. Each question was read by the examiner. If E felt that S did not understand a question or had misinterpreted it, he would repeat it. Each child was given as much time as he needed to answer the question.

All the special and regular class students were administered the Kohs learning potential procedure. The data were coded and anlayzed by chi-square procedures. There were two contrasts -- special versus regular class, and within the special class, by learning potential status. Learning potential data were available for the regular class sample as well, but there were few differences among these students when they grouped by learning potential status, and this data has not been reported.

Summaries of results

1. Social Interests and Activities.

There were few differences reported in the social interests and activities of these two samples. Their activities and interests lacked variety, and except for athletic interests, tended to be unstructured and focused on an interpersonal, belonging dimension rather than knowledge-oriented. The special class students tend to report themselves as more socially isolated and as peripheral group members.

The more able special class students by the learning potential assessment reported themselves to be more isolated socially, engaged in more passive activities, or in athletics, did not belong to peer groups, disliked group activities, and said they did not desire to change their situation. The less able (nongainer) students reported more active social involvements with their peers. But data from this and other studies indicates that nongainers give socially desirable responses which do not reflect their actual behaviors (Folman and Budoff, 1972b).

2. Family status

There were few differences between the two samples. Special class students tended to report they spent their free time with their families rather than friends. Both groups reported they were given responsible roles at home and the regular class students tended to report more responsibility in the home.

Nongainers reported themselves most alienated from their parents, and desired increased physical contacts, though not verbal interactions. High scorers, and gainers to a lesser degree, reported spending their free time outside the family, though they reported good relations with their families. High scorers reported having good relations with their fathers. Gainers reported good relations with their mothers, and desired better relations with their fathers (Folman and Budoff, 1972c).

3. Vocational Interests and Aspirations

a. Present school employment

(1) Special and regular class comparisons. Similarity of S's responses, more EMR Ss held previous job ($p < .07$) [they were one-year older, however]. More regular than EMR Ss obtained job on their own ($p < .08$).

(2) Learning Potential status (LP) effects within special class: (Nongainers, Gainers and High Scorers in that order) showed positive relation to: proportion of Ss who held jobs ($p < .10$); number of jobs held; length of employment; more NGs than Gs or HSs liked their job in nonspecific terms; more NGs than Gs and HSs got job on their own ($p < .06$); Gs wanted higher level jobs than NGs and HSs.

b. Vocational aspirations and expectations

(1) Special versus regular class: all Ss aspired to job similar to expectations. Special class Ss desired ($p < .05$) and expected lower level jobs ($p < .10$).

(2) LP comparisons -- More Gs and HSs aspired to lower level, blue collar jobs than NGs. NGs tended towards unrealistic aspirations but lowered this when asked expected job, to general level of special class students.

c. Level of understanding of vocational choice

(1) Special and regular class comparisons - 72% and 69% of special and regular class samples respectively indicated a degree of understanding along with job choice.

(2) LP comparisons -- More HSs gave thought to and had familiarity with proposed job and said they discussed the job with their family and others.

Discussion:

Special class members were able to get and hold jobs just as any other adolescent. Their lack of initiative in obtaining their own jobs may be due to stigmatized state. Regular class students had realistically higher job aspirations and expectations than the EMRs. Within the special class sample, findings supported the LP hypothesis that EMRs, classified as such by a psychometric criterion, are not a homogeneous group. The high scorer aspires to low level jobs, and displays a behavioral pattern consistent with minimizing risk of failure, and unlike the nongainers, displays little involvement in present low skill jobs. The nongainer behaves unrealistically and can be said to have retarded vocational development, supporting the LP assessment of nongainers as intellectually mentally retarded. IQ differences did not result in differences in realism of vocational choice, and the study emphasizes the importance of controlling for social status differences between EMR and nonEMR groups. A failure to do so accentuates the differences between "retardates" and "normals" while such a control diminishes differences in factors not related to academic success (Folman and Budoff, 1971).

d. Attitudes toward school and schooling.

The results indicated that:

(1) fewer special than regular class students saw a relationship between their present schooling and their future lives;

(2) while the majority of both groups aspired to finish high school and continue their education, more regular class Ss expected to do so;

(3) more regular than special class Ss tended to take personal responsibility for their failures in hypothetical locus of control questions, while special class Ss tended to blame themselves for their actual school failures;

(4) while most of the adolescents in both groups saw themselves as equal to or poorer students than their siblings, CA-peers, friends, and classmates, more special class Ss saw themselves as better academically than their friends and classmates;

(5) special class Ss tended to say they expend much effort in their school work and saw their school work as being the best they can do while the regular class students expressed more lackadaisical attitudes toward their school work.

Learning potential status within the special class sample was related to these academic variables. Like their regular class peers, the more able learning potential students (a) related school to their future adult job situation; (b) exhibited less discrepancy between their academic aspirations and expectations; (c) reported being given more responsible roles in hypothetical classroom situations; (d) exhibited an internal locus of control in both success and failure situations; and (e) reported putting more effort in their school work. Three were no differences in the special class students' self-perceptions of their school ability.

Discussion:

It must not be imagined, however, that the child segregated into special class and not thereby participating in the general stream of life in his junior high or high school is immune to the stigmatizing effects of this placement, no matter how devalued school may be in the social and value context of his community. While our findings indicate few major differences in the responses of these special and regular class students from a low income inner-city white community, the most able (LP) special class students do indicate some distress in their lives -- they tend to isolate themselves or may be isolated socially except for athletic activities; they aspire low vocationally though they desire more success in school and brighter prospects as adults. There is evidence in these data, in short, that they do feel stigmatized, "different," and have difficulty coping with it.

These data do offer further support for the broad range of differences within the psychometrically homogeneous special class population, and evidence in support of the hypothesis that the more able (LP) students' perceptions of themselves do tend to agree with those of their regular class peers, when both samples are drawn from the same low socioeconomic areas.

G. Learning Potential and Mother's Perceptions of her Child's Placement in an EMR Class.

Another source of data regarding the competence of these children outside school are the perceptions of the children's parents. In a pilot study, the mothers of 27 males who had been placed in special class and who had been examined with the learning potential procedure were interviewed in their homes. The sample was unbalanced -- 20 of the children were gainers, only seven were nongainers.

Three-fourths of the gainers' mothers felt their children learned less, and were slower than their peers. However, six of the twenty felt the reasons for the learning problems were "environmental" or external to the child and family; five felt the problem was only in special areas of study. But, they did not relate the child's failure in school to his success in life (steady job, happy marriage, etc.). Three-fourths of the gainers' mothers felt their children would "outgrow" their learning problems, and only one felt her child would not. Nineteen of the twenty felt their children would marry (nine had negative feelings about it; seventeen -- positive to positive-conditional reactions; three -- a sense of loss). This contrasts with nongainers' mothers -- four of the seven felt their children would marry, but only one had positive feelings about it. Only three of the twenty gainers had one or more siblings in special classes; five of the seven nongainers did.

The emotional reaction of mothers of gainers and nongainers to their child's placement in special class differed. Eleven gainers' mothers were surprised; twelve, upset at the time of placement (eleven were still upset by the time of the interview two years later). Six of the twelve felt the placement inappropriate, the other six were concerned about the stigma or possible connections with physical handicaps or mental illness. Of the nongainers' mothers, five of the seven said the placement "confirmed their suspicions." None said they felt the placement inappropriate.

Several kinds of questions were asked in order to ascertain whether mothers felt differently about, or treated their "EMR" child differently. Half of the gainers' mothers had "special worries" for their child, but at least on the surface these appeared to be the usual worries about an adolescent and his future. Gainers' and nongainers' mothers differed in the physical description of their children. Eleven gainers' mothers described their children as "average" or "different in a positive way." Four nongainers' mothers described their children as "different in a negative way," and only one said "different in a positive way." Four of the seven nongainers' mothers wanted play supervision for their children; most of the gainers' mothers did not. Informal peer group play was more successfully achieved by gainers, sixteen of which saw friends daily. Four nongainers saw their friends rarely or had none.

H. Learning Potential and Post School Adjustment.

The most powerful support for the Learning Potential argument can be drawn from data which indicate that the more able (LP) students from the special classes will attain economic and social independence as adults to a greater degree than the nongainers. During the past three years, students, 20 years and older, on whom we had collected Kohs Learning Potential data were located when possible. Only a small sample could be located for interviewers in part because 50% were no longer in their original communities and two school systems with large segments of the eligible sample permitted only mail questionnaires to be sent. The data presented summarizes the findings only with the interviewed sample.

A. Learning Potential and Post School Adjustment of Ex-Community Special Class EMRs. (Budoff).

1. Relationship between LP and vocational adjustment. An interview was held with 37 ex-special class young adults (CA 20 years), who had been tested on the Kohs LP procedure while they were in school

Learning Potential status: There was a significant relationship between learning potential status and military service classification. In the Boston area (but not in a small towns sample in Western Massachusetts), more able students by the LP criterion were more likely than the low able Ss, by this criterion, to have a draft status which implied mental and physical eligibility for military duty ($X^2 = 4.91$, $p < .05$). There was also a correlation between learning potential status and high school employment. Of those Ss who worked during the high school years, the more able (LP) students tended to have worked for a longer period of time (r = .42, p = .02). They also tended to have been drafted or enlisted in the armed services or to have been working full time more than 75% of the time since leaving school.

Initial administration of Kohs LP Task (K_1)*: Score on K_1 correlated with "reason for not seeking post-school training" and with "reasons for leaving last job." Ss with higher K_1 scores were more likely to report that they were unaware of any opportunities (as opposed to being disinterested) as a reason for not taking post-school training (r = -.48, p = .02). Ss with higher K_1 scores were also more likely to have left their last job for employer-initiated reasons as opposed to self-initiated reasons than Ss with lower K_1 scores (-.43, $p < .02$).

2. Relationship between LP and post-school social adjustment.

Learning Potential status: The higher the LP category, the more likely S was to report having a special group of friends (r = .31, p = .04). Ss in the higher LP categories also tended to report a higher rate of participation in high school groups (r = .32, p = .02) and were more likely to report that they dated (r = .43, p = .02).

Immediate post training LP scores (K_2)*: Higher K_2 scorers were more likely to live away from home (r = .31, p = .03) and were more likely to have a drivers' license (r = .39, p = .01).

B. Learning Potential and Institutional Discharge Status Among Young Adult School-Age Defined Educable Mental Retardates (Budoff, M.)

The institutional status of the students in the original studies with Kohs Learning Potential Test (Budoff and Friedman, 1964; Budoff, 1967) from 7-10 years following the original data collections was examined. The hypothesis was that the more able (LP) students would no longer be in residence. The less able (LP) students' ability to profit from experience on the LP test would still be in residence at the institution, having been independently judged by the staff as less capable of establishing economic and social independence.

A survey of the discharge status of institutionalized educable mentally retarded patients who had been studied with the Kohs Learning Potential (LP) assessment measure in the early 1960's was undertaken. Learning Potential was strongly associated with the fact of discharge. The more able patients by the Learning Potential criterion also had higher scores on various performance ability measures but these higher scores were not systematically related to discharge status. Among patients institutionalized below 11 years of age, high Learning Potential status Ss had been discharged, whereas less able (LP) patients were not.

LP status does appear to relate to subsequent adequacy of adult adjustment.

Summary of the construct validity for learning potential Assessment.

At the start of this project, the data we have accumulated indicate a broad range of consistency and support for the differences shown by psychometrically defined educable mental retardates (EMRs) to this training-based assessment procedure. More able children, by the learning potential criterion, indicate areas of more adequate functioning when performance does not depend on satisfactory progress in school. Those children who have demonstrated their competence on the LP task (e.g., the Kohs Block Designs) also tended to score higher on other nonverbal or minimally verbal tasks, such as the Wechsler Intelligence Scale for Children Performance IQ (PIQ), and Raven Progressive Matrices. The personality data, whether behavioral or verbal, suggests a greater sense of personal adequacy in the more able children, greater flexibility on a concept shift task, more flexibility and ability to delay response, ability to manage frustrations more positively, a view of oneself as less neurotic or maladjusted, the tendency to set aspiration levels consonant with one's performance even under the stress of failure, and the feeling that one is not manifestly different socially from one's "nonretarded" peers. However, they do realistically feel inadequate in academic areas (Folman & Budoff, 1972). The students who seem to be most able by the learning potential criterion are also those who seem to be making the more concrete and realistic vocational plans. The competence and understanding of the more able children can more generally be demonstrated by their differential ability to build empirically meaningful relationships among objects and/or events, when they are not required to state verbally the abstract elements of their empirical understanding.

The more able (LP) students' ability to perform in this manner was clearly demonstrated in a study in which they were taught electricity empirically by manipulating batteries, bulbs, and wires. Following teaching, they were able to display understanding of the principles of electricity on a nonverbal task to the same degree as low achieving higher IQ children who were taught the curriculum. In these instances, they performed as well as average achieving children. The performance of special class children seems to indicate their difficulties are in the verbal expression of their understanding (Budoff, Meskin, & Harrison, 1971; Harrison & Budoff, 1972; Suzman, 1973; Hutten, 1973; McCaffrey, 1973). For example, Suzman, (1973) who allowed EMRs to paraphrase embedded sentences verbally or in mime, reported that the difficulty of special class children appears to be in production of verbal responses, not in the understanding of the meaning of the sentence. They were able to paraphrase the message conveyed in the embedded sentence when the paraphrase was performed in mime, with few utterances required. Mandelkorn (1972) and Mandelkorn & Corman (1973) have shown that the more able students, by the LP criterion, profited from a short training unit which taught verbal communication skills to a greater degree than the less able special class students.

The most tenable hypothesis by which to explain these consistent differences between more and less able special class students is to consider the high able group by the learning potential criterion (high scorers and gainers) as educationally handicapped. These are children who have been unable to progress satisfactorily in school for whatever reasons, but who seem to have ability not usually ascribed to them.

Those who do not profit from the experience of training on the nonverbal reasoning tasks (nongainers) seem to function on a wide variety of tasks as mentally retarded children are described. The high able group by the LP criterion may be those children who have been designated by some as "pseudo-feebleminded." These are school-aged children diagnosed as mentally retarded. As adults in the wider opportunity system of the society, they become economically and socially independent, confounding the previous diagnosis which guided the educational opportunities made available to them. As a result, these children leave school as functionally illiterate. For a large proportion of EMR children (about 2/3 of those noninstitutionalized and over 15 years) the learning potential procedure evoked a positive response and indications of previously unsuspected ability to reason. This proportion, coincidentally, is the approximate proportion of school-aged IQ-defined "EMRs" who attain an independent social and economic adjustment during their adult years.

This evidence of ability to learn empirically should remove the more able students by the LP criterion from consideration as mentally retarded, though they seem to suffer from serious verbal-expressive difficulties which hamper their educational progress. In at least one instance, these verbal problems may be amenable to training (Mandelkorn & Corman, 1973).

In the following section, full length reports of the studies performed on this project will be presented.

PART II

Studies Performed on This Project
Grant NE-G-00-3-0016

Study 1

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure. *RIEP-Print* No. 56, 1973. [ED 086 924]

Study 2

Budoff, M., & Corman, L. The effectiveness of a group training procedure on the Raven learning potential measure with children from diverse racial and socioeconomic backgrounds. *RIEP-Print* No. 58, 1973. [ED 086 923]

Study 3

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure with black and white special class students. *RIEP-Print* No. 57, 1973. [ED 085 969]

Study 4

Corman, L., & Budoff, M. IQ and learning potential measurements of general intelligence: A comparison of relationships. In D. A. A. Primrose (Ed.), *Proceedings of the Third Congress of the International Asssociation for the Scientific Study of Mental Deficiency* (Vol. I). Stirlingshire, Scotland: IASSMD, 1975.

Study 5

Corman, L., & Budoff, M. Factorial composition of Raven's Progressive Matrices in the context of learning potential assessment. *RIEP-Print* No. 62, 1973/1975.

Study 6

Budoff, M., & Corman, L. Factor analysis of error types in Raven Progressive Matrices: A failure to obtain logical structure. *RIEP-Print* No. 79, 1975.

Study 7

Budoff, M., & Corman, L. Effectiveness of a learning potential procedure in improving problem-solving skills of retarded and nonretarded children. *RIEP-Print* No. 63, 1975.

Study 8

Corman, L., & Budoff, M. The picture word game: a nonverbal test of the ability to use language-related symbols. *RIEP-Print* No. 77, 1974. [ED 108 439]

Study 9

Budoff, M. A further study of the educational validity of learning potential assessment: Response to objectives-based curricula. 1976.

Study 10

Corman, L., & Budoff, M. Factor structures of Spanish-speaking and non-Spanish-speaking children on Raven's Progressive Matrices. *Journal of Educational and Psychological Measurement*, 1974, *34*, 977-981.

Study 11

Budoff, M., Corman, L., & Giman, A. An educational test of learning potential assessment with Spanish-speaking youth. *Interamerican Journal of Psychology*, 1976, *10*, 13-24. [ED 108 436]

Study 12

Budoff, M., & Hamilton, J. L. Optimizing test performance of the moderately and severely mentally retarded. *American Journal of Mental Deficiency*, 1976, *81*, 49-57.

Study 13

Babad, E., & Bashi, J. Final Report: An educational test of the validity of learning potential measurement. Ford Foundation, June, 1975, *RIEP-Print* No. 91, 1975.

PLEASE REFER TO: ED 086 924

STUDIES IN LEARNING POTENTIAL

A COMPARISON OF GROUP AND INDIVIDUAL TRAINING PROCEDURES ON THE RAVEN LEARNING POTENTIAL MEASURE

by

Louise Corman and Milton Budoff

Volume 3, Number 56

1973

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

PLEASE REFER TO ED 086 924

A COMPARISON OF GROUP AND INDIVIDUAL TRAINING PROCEDURES ON THE RAVEN LEARNING POTENTIAL MEASURE[1]

Louise Corman and Milton Budoff

Research Institute for Educational Problems[2]

The purpose of the study was to determine whether a training procedure involving groups of 12 to 15 students was as effective as the individual training procedure used in the past with respect to improving performance on the Raven Learning Potential measure. Since the Raven Progressive Matrices has been found to be an effective learning potential measure when students received individual training (Budoff, 1972), the development of a group training procedure was considered important in increasing the cost-effectiveness of the LP procedure. An increase in cost-effectiveness, however, would be valuable only if it was not attained at the expense of the performance capabilities evidenced by these children of concern.

Method

The sample in the study consisted of 202 first through fifth grade students in an urban community in Massachusetts. The sample was almost entirely white and heterogeneous with regard to social class. Two classes from each of the five grade levels were selected from one elementary school. At the beginning of the study, Sets A, AB, B of the Raven Progressive Matrices were group administered to students in each of the ten classrooms, one class at a time. At this time each student's age and sex were recorded, and his father's occupation was obtained from school records.

On the basis of scores attained at this test administration, each student was assigned to one of three groups: a group which received group training on the Raven test, a group which received individual training, and a group which received no training and served as a control group.

The assignment procedure for all students in the two classrooms within each grade was as follows: Raven scores of all students in any one grade were rank ordered from low to high. In the event of tied scores a rank position for the two scores was randomly assigned. The three students with the three highest scores constituted a block, and each student within this block was randomly assigned to one of the three groups. Each of the three students with the three lowest scores was then randomly assigned to one of the three groups. This procedure was repeated for each block of three students with the next highest and next lowest scores, working toward the middle of the score distribution.

Students from the two classrooms in each grade who were assigned to the group training group were trained on the Raven Learning Potential procedure together, so that each group training session involved 12 to 15 students.

During the week after group and individual training had taken place, all students were posttested. First graders received Sets A, AB, B. Because initial mean scores became progressively higher at each grade level (17.2, 20.7, 22.4, 26.3, and 27.6 for grades 1 to 5, respectively), it was decided that a number of items on Set C of the Raven test should be included on the posttest for children in the second through fifth grade. In addition to the 36 items in Sets A, AB, B, second graders received the first three items of C, third graders received the first six, fourth graders received the first nine, and fifth graders received all twelve. The maximum possible posttest score was therefore 36 for first graders, 39 for second graders, 42 for third graders, 45 for fourth graders, and 48 for fifth graders. By reducing the likelihood of a ceiling effect on the posttest in this way, it was hoped that students in all grades would have the opportunity to demonstrate improved performance on the posttest. The interval between pretesting and posttesting was approximately three weeks.

A stepwise multiple regression equation was performed with posttest score (R2) as the dependent variable. Six independent variables were entered into the equation in the following order: (a) pretest score (R1), (b) age, (c) sex (coded 1 = male, 2 = female), (d) rating of father's occupation on the Turner Scale (1964), (e) Group 1 =

membership in either training group (coded 1) versus membership in the control group (coded 2), and (f) Group 2 = membership in the group trained group (coded 1), the control group (coded 2), or the individually trained group (coded 3). The latter two variables represented orthogonal contrasts of the training group factor. Partial correlation coefficients of two-way interactions involving these six variables, when the six effects had been entered into the equation, were also obtained.

Results

Means and standard deviations of the three groups on the pretest and the posttest are presented in Table 1. The table indicates that the blocking procedure for assigning students to groups was highly effective in equalizing both the initial means and variances of the three groups. It is evident that the control group improved to some extent on the posttest, probably as a result of practice in taking the test. This group raised their score an average of 4.65 points from pretest to posttest. Both the individually trained and group trained groups demonstrated a higher mean increase than that of the control group: the mean of the group trained group increased eight points, and the mean of the individually trained group increased nine points from pretest to posttest. The mean increases of the two trained groups, then, were similar to each other and greater than the average gain (4.65) of the control group.

Insert Table 1 about here

Table 2 presents the results of the stepwise multiple regression equation on posttest scores. Pretest score (Rl), age, and Group 1 (membership in either trained group versus membership in the control group) were all significantly related to posttest scores. The negative sign of the beta weight of the Group 1 factor indicated that subjects in either training group performed significantly better on the posttest than students in the control group. The fact that the Group 2 factor was not significant ($\underline{p}$ = .217) indicated that there was no difference in posttest scores of students who received group training and students in the individually trained group. In addition, Raven posttest scores did not differ for students from different socioeconomic backgrounds, nor did they differ between males and females. The percent of variance accounted for by all variables in the equation was 71.6; the multiple $\underline{r}^2$ was .716 ($\underline{F}$ = 81.60, $\underline{df}$ = 6/194, $\underline{p}$ <.001).

Inspection of the partial correlation coefficients between posttest scores and the 14 two-way interactions, after main effects had all been entered into the equation, indicated that only one was significant. The partial correlation coefficient of Rl X Group 2 was -.166 ($\underline{p}$ <.05).

TABLE 1

Means and Standard Deviations of Three Groups on Raven Pretests and Posttests

	Pretest		Posttest		
Group	$\bar{X}$	SD	$\bar{X}$	SD	N
Group trained	22.882	6.502	30.750	8.819	68
Control	22.866	6.517	27.522	9.437	67
Individually trained	22.746	6.527	31.727	8.229	67

TABLE 2

Results of Multiple Regression on Raven Posttest Scores

Variable	Beta	T-test	df
R1	.581	12.12*	194
Age	.348	7.71*	194
Sex	.067	1.74	194
SES	.020	0.47	194
Group 1 (trained vs. control)	-.203	-5.27*	194
Group 2 (group, control, or individual)	.047	1.24	194

$r^2 = .716$

$F = 81.60$, $df = 6/194$*

*$p < .001$

Graphs of this interaction indicated that students with low pretest scores who received individual training got higher posttest scores than students with low pretest scores who received group training.

Discussion

The study has demonstrated that a group training procedure may be used with the Raven LP measure, and that, on the whole, results of group training with intellectually normal children are similar to those that are achieved when these children are individually trained. The significant interaction between training group and pretest score might suggest that the individual training procedure may be more effective with lower IQ children, e.g., educable mental retardates, than a group training procedure, since these children are more likely than normal children to have low pretest scores.

References

Budoff, M. Measuring learning potential: An alternative to the traditional intelligence test. *Studies in Learning Potential*, 1972, *3*, No. 39.

Turner, R. H. *Social context of ambition: A study of high school seniors in Los Angeles*. San Francisco: Chandler Publications, 1964.

Footnotes

[1] This research was supported by Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education and Grant NE-G-00-3-0016 from the National Institute of Education, both under the Department of Health, Education, and Welfare. We extend our appreciation for cooperation in the conduct of this study to Mr. Charles Burack, Principal, Elm Park Community School, Worcester, Massachusetts, and Mr. Evans Tsoulas, who coordinated and expedited the data collection effort.

[2] 12 Maple Avenue, Cambridge, Massachusetts 02139.

PLEASE REFER TO ED 086 923

STUDIES IN LEARNING POTENTIAL

THE EFFECTIVENESS OF A GROUP TRAINING PROCEDURE ON THE RAVEN LEARNING POTENTIAL MEASURE WITH CHILDREN FROM DIVERSE RACIAL AND SOCIOECONOMIC BACKGROUNDS

by

Milton Budoff and Louise Corman

Volume 3, Number 58

1973

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

PLEASE REFER TO ED 086 923

THE EFFECTIVENESS OF A GROUP TRAINING PROCEDURE ON THE RAVEN LEARNING POTENTIAL MEASURE WITH CHILDREN FROM DIVERSE RACIAL AND SOCIOECONOMIC BACKGROUNDS

Abstract

This study determined the effectiveness of a group training procedure in improving performance of intellectually average students on the Raven Learning Potential measure, and examined the differential effectiveness of training by race and social class. The sample of 379 second through fifth graders, half of whom were black, were pretested on Raven's Coloured Progressive Matrices, randomly assigned to trained or nontrained groups, and posttested on the Raven test. Multiple regression analysis on posttest scores indicated that trained students improved their nonverbal reasoning ability more than nontrained students and that training was equally effective with black and white children. Middle class children were found to benefit more from training than lower class children.

THE EFFECTIVENESS OF A GROUP TRAINING PROCEDURE ON THE RAVEN LEARNING POTENTIAL MEASURE WITH CHILDREN FROM DIVERSE RACIAL AND SOCIOECONOMIC BACKGROUNDS[1]

Milton Budoff and Louise Corman
Research Institute for Educational Problems

The purpose of the study was to determine whether a group training procedure was effective in improving performance on the Raven Learning Potential (LP) measure by intellectually average students from heterogeneous socioeconomic and racial backgrounds. A previous investigation (Corman & Budoff, 1973a) with 202 first through fifth grade white students demonstrated that training these students in groups was as effective in improving their performance on the Raven measure as the individual training which had been given in prior studies of learning potential. On the basis of these findings, a group training procedure was employed in the present study and the possible differential effectiveness of this procedure by race or social class was examined.

Method

The sample consisted of 379 first through fifth grade students in an urban area of upstate New York. Four classes from each of the five grade levels were selected from two elementary schools, one in a low socioeconomic district and one in a middle class community which bussed in large numbers of black students. Fifty-three percent of the sample was black. Twenty-eight percent of the students' fathers were employed in business or professional occupations; the majority were blue collar or clerical workers. Intelligence test scores were not available for these students; mean scores on standardized achievement tests, however, indicated that students in all grades were performing below grade level in reading and math, and that the discrepancy between average performance and national norms became progressively greater from second to fifth grade.

At the beginning of the study, Sets A, AB, B of the Raven Progressive Matrices were group administered to students in each classroom, one class at a time with the standard instructions (Raven, 1965). At this time each student's sex and race were recorded, and his father's occupation, birthdate, and achievement test scores were obtained from school records. On the basis of scores attained at this test administration, each student was assigned to one of two groups: a group which received training on the Raven test or a group which received no training and served as a control group.

The assignment procedure for all students in each classroom was as follows: Raven scores of all students in any one class were rank ordered from low to high. In the event of tied scores a rank position for the two scores was randomly assigned. The two students with the two lowest scores constituted a block, and each student within this block was randomly assigned to one of the two groups. Each of the two students with the two next lowest scores was then randomly assigned to one of the two groups. This procedure was repeated for each block of two students until all students had been assigned to the two groups. Students in one classroom who were assigned to the training group were trained together, so that training groups originally contained nine to thirteen students. Because of absenteeism, however, actual group sizes ranged from six to thirteen; students who were absent from either training session were deleted from the sample in an effort to avoid distortion of the random assignment procedure.

All students were trained and posttested on Sets A, AB, B within two weeks after pretests had been given. In order to allow older students to demonstrate improvement following training (i.e., to remove a possible ceiling effect on the test), third through fifth grade students were given Raven Sets C and D in addition to Sets A, AB, B on the posttest. Maximum possible scores, then, were 36 on the pretest for all students, 36 on the posttest for first and second graders, and 60 on the posttest for third through fifth graders.

A training booklet which contained nontest problems dealing with pattern completion, orientation of elements within a pattern and double classification problems was distributed to each child. The trainer presented the problems on 2 X 2 slides from a Kodak Carousel projector on a blackboard. The students were required to draw in the missing element for the design and mark the appropriate choice from the six choices presented on the lower half of the page. For the double classification problems, it was found that children could easily derive one attribute at a time but often did not hold the first attribute in mind while they derived the second relevant attribute. During development of the training procedure, the child's understanding was facilitated by having him draw the relevant attributes, one at a time, as he derived them. This helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems mentally with very little trouble. The requirements of each problem type were presented in meaningful designs initially, e.g., an American flag with a piece missing, and then a geometric form to attune the child to the basic format of the Matrices test problems. Individual children were called to indicate the correct choice, and to give reasons for their choice. A slide with the answer included allowed them to compare their choice and to correct it, if necessary. Three scores are derived--pre- and posttraining score, and posttraining score adjusted for pretraining score.

A stepwise multiple regression equation was performed with posttest score (R2) as the dependent variable. Six independent variables were entered into the equation in the following order: (a) pretest score (R1), (b) Age, (c) Sex (coded 1 = male, 2 = female), (d) SES rating of father's occupation on the Turner (1964) scale (coded 0 = on welfare to 9 = managerial), (e) Race (coded 1 = black, 2 = white), (f) Training Group (trained coded 1 or control coded 2). Partial correlation coefficients of two-way interactions involving these variables, when the six effects had been entered into the equation, were also obtained.

Results

Pretest and posttest means and standard deviations of the two groups in each grade are presented in Table 1. The table indicates that the randomized blocking procedure for assigning students to groups was highly effective in equalizing the initial means and variances of the groups. In the total sample, the mean of the control group rose 8.2 points from pretest to posttest, probably as a result of practice in taking the pretest, while the mean increase of the trained group was 10.9 points. The steady rise of pretest scores from first to fifth grade suggests that ability to do the Raven problems increases with age. The marked difference in means on the total posttest between second and third grades reflects the fact that third through fifth graders received the 60 item test.

TABLE 1

Means and Standard Deviations of the Two Groups by Grade on Raven Pretest and Posttest

		R1 (A,AB,B)		R2[a]		R2 (A,AB,B)		
Grade	Group	X̄	SD	X̄	SD	X̄	SD	N
1	trained	16.73	5.87	21.24	6.69	same as R2		33
	control	16.76	6.22	18.59	6.08	same as R2		41
2	trained	20.03	6.62	24.46	6.89	same as R2		35
	control	20.90	5.78	21.38	6.39	same as R2		39
3	trained	23.58	5.45	37.25	8.94	26.39	5.53	36
	control	24.14	5.16	36.20	8.96	26.07	4.87	44
4	trained	26.03	6.63	41.65	11.50	28.08	6.83	40
	control	25.46	5.90	39.20	12.06	27.30	6.75	46
5	trained	26.90	6.32	43.57	8.52	30.10	4.35	29
	control	26.94	4.40	41.00	8.26	27.97	4.06	36
Total	trained	22.68	7.16	33.63	12.48	26.03	6.78	173
sample	control	22.84	6.56	31.08	12.67	24.21	6.76	206

[a] Grades 3, 4, and 5 were given Sets C and D in the post-training administration in addition to Sets A, AB, B.

Table 2 presents the results of the stepwise multiple regression equation on posttest scores. Pretest score (R1), Age, SES, and Training Group were all significantly related to posttest scores ($p < .001$). The negative sign of the beta weight of the Training Group factor indicated that subjects in the trained group performed significantly better on the posttest than students in the control group. Older students or students from middle class backgrounds attained higher posttest scores than younger students or students from lower SES backgrounds. Posttest scores were not differentially affected by race or sex. The percent of variance accounted for by all variables in the equation was 73.8, with 20.8% attributable to pretest score; the multiple r^2 was .738 ($F = 170.79$, 6/364 df, $p < .001$).

Inspection of the partial correlation coefficients between posttest scores and the 15 two-way interactions, after main effects had all been entered into the equation, indicated that the following interactions were significant: R1 X Age ($r = .198$, $p < .01$), R1 X SES ($r = .148$, $p < .01$), Age X SES ($r = .209$, $p < .01$), and SES X Training Group ($r = .123$, $p < .05$).

Plots of these interactions provided supplementary information to that revealed by the significant main effects: middle class students attained higher posttest scores, particularly if they had high pretest scores or were older. The significant R1 X Age interaction indicated that older

TABLE 2

Results of Multiple Regression on Raven Posttest Scores

Variable	Beta	T-test	*df*
R1	.592	16.99*	364
Age	.350	11.23*	364
Sex	.028	1.02	364
SES	.114	3.73*	364
Race	.016	0.57	364
Training group	-.095	-3.52*	364

$r^2 = .738$

$F = 170.79$, *df* = 6/364*

*p <.001

students with high pretest scores got high scores on the posttest.

The SES X Training Group interaction was of particular interest. Table 3 presents the Raven pretest and posttest means for trained and nontrained subjects in three social class groups, when social class ratings on the Turner Scale (1964) were trichotomized into low (welfare, unskilled, or semi-skilled), middle (skilled or clerical), and high (business or professional) levels. The table indicates that the discrepancy in posttest mean gains between the trained and nontrained groups was greatest in the middle and high SES ranges. There was almost no difference in posttest mean gains between the trained and nontrained low SES groups. Even without training, the low SES students increased their mean scores to the levels of their trained low and middle income peers.

Discussion

Perhaps the most important findings of this study were that the training procedure was successful in improving students' nonverbal reasoning ability as measured by the Raven Progressive Matrices, and that the training was equally effective with black and white students. Sixty-five percent of the black students were from low SES backgrounds, compared to 49% of the white students in the sample. The partial r for the Race X Training interaction on posttest

TABLE 3

Mean Raven Scores of Trained and Nontrained Subjects in Three Social Class Groups

		Trained			Nontrained		
		$\bar{X}$	SD	N	$\bar{X}$	SD	N
Low SES	R1	20.41	6.75	63	21.85	6.44	88
	R2	29.60	11.95	62	30.47	11.90	86
	difference	9.19			8.62		
Mid SES	R1	21.32	6.62	19	22.50	6.35	20
	R2	30.68	13.62	19	28.32	12.98	19
	difference	9.36			5.82		
High SES	R1	28.03	6.56	37	27.32	5.21	37
	R2	42.11	10.93	37	39.06	11.74	36
	difference	14.08			11.74		

scores was not significant before SES was entered into the equation. Therefore, the lack of differential training effects by race was not the result of a confound between race and social class. The finding that males often attained higher Raven scores that has emerged in previous studies with special class students (Corman & Budoff, 1973b; Corman & Budoff, 1973c) was not found with this sample.

The significant social class effect and SES X Training interaction obtained with this sample were not obtained in a prior investigation (Corman & Budoff, 1973a) which compared the relative effectiveness of group and individual training procedures with students in the same age range as this sample. While the prior sample consisted almost entirely of white students, their mean and standard deviation on the Turner Social Classification of Occupations Scale was very similar to the sample in this study. Despite this similarity in socioeconomic composition of the samples in the two studies, the previous investigation found training to be equally effective with students of different SES levels. With the present sample, however, training was found to be more beneficial to students from middle and high SES backgrounds.

One may more meaningfully assess the impact of the training by examining the relative standing of the lower SES students prior to and following training vis-a-vis the pretest scores of the middle class children. The premise is that the pretest levels of the middle class child represent a criterion for good school and test performance. Learning potential assessment compensates for the middle-class children's lifelong spontaneous acquisition by providing the low income child by providing experiences that equip him to deal more adequately with the reasoning task at hand. One intent of a training-based assessment is to diminish the competence gap between more and less privileged groups.

The investigators set the criterion for effectiveness of the training at the proportion of low income students who attain the pretest level of their more privileged agemates following training. Lower class was defined as 0 to 4 on the Turner Scale (welfare, unskilled, skilled blue collar, or clerical), and middle class as 5 to 9 (business and professional). On the pretest, only 15% of the lower class students reached or exceeded the pretest mean of the middle class students at their age level; on the posttest, however, 40% of the lower class students attained scores as high or higher than the middle class pretest mean.

Babad and Budoff (in press) and Corman and Budoff (1973d) reported similar findings with a different nonverbal reasoning task (a series test) which was administered in the test-train-retest paradigm. Babad and Budoff (in press) included samples of students from special classes for the educable mentally retarded (IQ < 80), and those with IQs in the 80 to 100, and 100 plus ranges. Sixteen percent of the educable mentally retarded sample fell at or above the pretest means of the dull normal sample (mean IQ difference was 17 points), and 36% exceeded this pretest mean after training. With the bright normal group as criterion for the EMRs (there was a 45-point difference between the mean IQs of the two groups), none of the EMRs reached the bright normal mean in the initial test; 13% exceeded it following training. Eleven and 35% of the dull normal group attained the pretraining mean of the bright normal group prior to and following training, respectively.

Corman and Budoff (1973d) reported that 51.3% of their low income normal sample exceeded the middle class pretest mean. What is notable in this study is that the low SES sample was largely black (approximately 80%) and the contrast middle SES students lived in affluent white communities in western Connecticut.

Providing low SES students with experience with the problem-solving styles commonly acquired spontaneously by middle class students clearly minimized the usual SES effect for a significant proportion of these students, when the task is nonverbal reasoning problems. In the three studies this effect was minimized with less than one hour of group-administered training. Even when the comparisons are made with posttraining scores in the present study, the SES main effect only accounts for 1% of the variance. A learning potential assessment procedure which provides equalizing experience by including task-relevant training within the assessment process shows promise of being a culture fair means of assessing ability to profit from experience, i.e., intelligence.

References

Babad, E. Y., & Budoff, M. Sensitivity and validity of learning potential measurement in three levels of ability. Journal of Educational Psychology, in press.

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure. Studies in Learning Potential, 1973, 3, No. 56. (a)

Corman, L., & Budoff, M. IQ and learning potential measurements of general intelligence: A comparison of relationships. Proceedings of the Third Congress of the International Association for the Scientific Study of Mental Deficiency, in press. (b)

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measure with black and white special class students. Studies in Learning Potential, 1973, 3, No. 57. (c)

Corman, L., & Budoff, M. The Series Test as a measure of learning potential. Studies in Learning Potential, 1973, 3, No. 47. (d)

Raven, J. C. Guide to using Raven Progressive Matrices. London: H. K. Lewis, 1965.

Turner, R. H. Social context of ambition: A study of high school seniors in Los Angeles. San Francisco: Chandler Publications, 1964.

Footnotes

[1] This research was supported by Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education, and Grant NE-G-00-3-0016 from the National Institute of Education, both under the Department of Health, Education, and Welfare. The statistical consultation of Robert H. Harrison is gratefully acknowledged.

PLEASE REFER TO ED 085 969; EC 060 957

STUDIES IN LEARNING POTENTIAL

A COMPARISON OF GROUP AND INDIVIDUAL TRAINING PROCEDURES ON THE RAVEN LEARNING POTENTIAL MEASURE WITH BLACK AND WHITE SPECIAL CLASS STUDENTS

by

Louise Corman and Milton Budoff

Volume 3, Number 57
1973

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

PLEASE REFER TO ED 085 969; EC 060 957

A COMPARISON OF GROUP AND INDIVIDUAL TRAINING PROCEDURES ON THE RAVEN LEARNING POTENTIAL MEASURE WITH BLACK AND WHITE SPECIAL CLASS STUDENTS[1]

Louise Corman and Milton Budoff

Research Institute for Educational Problems[2]

The purpose of the study was to determine whether a training procedure involving groups of five to eight special class students was as effective as the individually administered learning potential (LP) procedure used in the past, with respect to improving performance on Raven's Progressive Matrices. Since Raven's Progressive Matrices has been found to be an effective learning potential measure with individually trained educable mentally retarded students (Budoff, 1972), the development of a group training procedure was considered important in increasing the cost-effectiveness of the LP procedure. A previous investigation with intellectually normal children (Corman & Budoff, 1973a) demonstrated that training in groups was as effective as individual training in improving the performance on the Raven measure by these students as a whole. The goal of this study was to replicate these findings with special class students.

Method

The sample consisted of 174 special class students in five schools of an urban community in upstate New York. The students ranged in age from 67 to 176 months, with a mean age of 128.6 months (SD = 23.6). All but 5% were from low socioeconomic backgrounds (welfare, blue collar, or clerical) as indicated by father's occupation, and 43.4% were black. The mean IQ of the sample on the WISC or Stanford-Binet was 68.4, with a standard deviation of 9.0.

Procedure

Sets A, AB, B of the Raven Progressive Matrices were administered to students in groups of approximately 15 to 25, depending on the number of special class students in each school. In schools which contained more than 25 special class students, students were tested in two groups of equal size. Each student's sex, race, father's occupation, birthdate, and IQ were recorded from school records. On the basis of Raven scores attained at this test administration, each student was assigned to one of three groups: those which received group training, individual training, and no training (control group).

The assignment procedure for all students in each group tested was as follows: Raven scores of all students were rank ordered from low to high. In the event of tied scores, a rank position for the two scores was randomly assigned. The three students with the three highest scores constituted a block, and each student within this block was randomly assigned to one of the three groups. Each of the three students with the three next highest scores was then randomly assigned to one of the three groups. This procedure was repeated for each block of three students until all students had been assigned to one of the three training groups. Since one third of the students in each group tested were assigned to the group trained group, each group training session involved five to eight students. All students were trained and posttested on Sets A, AB, B within a week after pretests had been given. White female testers and trainers were used throughout the study.

Training Procedure

Logical analysis of the test items indicates that pattern completion, orientation of elements, and analogy problems are readily identifiable types of problems in Sets A, AB, B. This observation was essentially confirmed by results of recent factor analyses of item responses of children in special and regular classes (Corman & Budoff, 1973b). A training booklet containing problems of these types was formulated and distributed to each child for use in two training sessions. The training procedure embodied the following practices and principles:

1. No problem which appears in Raven's Matrices was presented in the training booklet so that there was no specific training toward the actual test problems.

2. Great efforts were made to be certain that the child understood the demands made of him on the task. Hence, the training procedure was not a strictly standardized one but allowed for a number of variations to ascertain whether the child understood what was required before the conclusion was drawn that he was unable to master the particular type of item. Problems were presented in up to three different formats before training was discontinued.

3. The requirements of the task were demonstrated initially on materials that are pictorial and meaningful, i.e., incomplete American flag or a Campbell soup can for the pattern completion tasks. Successive problems in the training series required the child to deal with geometric forms in order to attune the child to the basic format of the matrices test presentation.

4. The requirements of the task were concretized in motoric performances, e.g., having the child draw the item that completes the pattern before perusing the various alternative solutions presented under the problems. In the double classification problems, it was found that children could easily derive one attribute at a time but often did not hold the first attribute in mind while they derived the second relevant attribute. During development of the training procedure, the child's understanding was facilitated by having him draw the relevant attributes, one at a time,

as he derived them. This helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems mentally with very little trouble.

It should be noted that when a child had difficulty drawing the designs, the trainer drew them for him.

5. During the training on the double classification problems, after the child ceased drawing the alternatives and began doing the problems mentally, the child was required to indicate verbally how he arrived at his solution on every third problem. This seemed to help the child maintain the quality of the solution process attained by the drawing and the trainer's accompanying verbalization. A minimal verbal response, often in combination with pointing responses by the child, sufficed to indicate to the trainer that the child was solving the problems.

Detailed presentation of the training procedure is available in Budoff, Corman, and Litzinger (1973).

Statistical Analyses

A stepwise multiple regression equation was performed with posttest score (R2) an the dependent variable. Seven independent variables were entered into the equation in the following order: (a) pretest score (R1); (b) Age; (c) Sex (coded 1 = male, 2 = female); (d) SES (rating of father's occupation on the Turner Scale (1964), coded from

0 = on welfare to 9 = managerial); (e) Race (coded 1 = black, 2 = white); (f) Training Condition = coached (membership in either training group, coded 1) versus noncoached (control, coded 2); and (g) Training group = membership in the group trained group (coded 1), the control group (coded 2), or the individually trained group (coded 3). The latter two variables represented orthogonal contrasts of the training group factor. Partial correlation coefficients of two-way interactions involving these seven variables, when the seven effects had been entered into the equation, were also obtained.

Results

Means and standard deviations of the three groups on the pretest and the posttest are presented in Table 1. The table indicates that the blocking procedure for assigning students to groups was highly effective in equalizing the initial means and variances of the three groups. It is evident that the control group improved slightly on the posttest, probably as a result of practice in taking the pretest. This group raised their score an average of 1.3 points from pretest to posttest. Both the individually trained and group trained groups demonstrated a higher mean increase than that of the control group: the mean of the group trained group increased 3.5 points, and the mean of the individually trained group increased 3.2 points from pretest to posttest. The mean increases of the two trained groups, then, were similar to each other and greater than the average gain of the control group.

Insert Table 1 about here

Table 2 presents the results of the stepwise multiple regression equation on posttest scores. The percent of variance accounted for by all variables in the equation was 69.6 with 36.9% attributable to pretest score; the multiple r^2 was .696 ($F = 51.34$, $df = 7/157$, $p < .001$). Pretest score (R1), Age, Sex, Race, and Training Condition (membership in either trained group versus membership in the control group) were all significantly related to posttest scores. These effects were significant at the .01 level or beyond with the exception of Race, which accounted for 1% of the unique variance in R2 and was significant at the .05 level. The negative sign of the beta weight of the Training Condition factor indicated that subjects in either training group performed significantly better on the posttest than students in the control group. The fact that the Training Group factor was not significant ($p = .299$) indicated that there was no difference in posttest scores of students who received group or individual training. Older students, males, or white children achieved higher posttest scores than younger students, females, or blacks. The main effect of social class was not a significant predictor of performance following training. This finding probably is due to the constricted SES range of the sample. Also, the usual finding is that middle class students in special classes often have various physical involvements not clinically evident in the low SES students.

Insert Table 2 about here

Inspection of the partial correlation coefficients between posttest scores and the 20 two-way linear by linear interactions, after main effects had all been entered into the equation, indicated that the following interactions were significant: (a) Sex X Training Group ($r = -.165$, $p < .05$), (b) Race X Training Group ($r = .160$, $p < .05$), (c) Social Class X Training Group ($r = .214$, $p < .01$), and (d) Social Class X Race ($r = -.188$, $p < .05$). Each of these interactions will be examined in turn in the discussion which follows.

Cell means in the Sex X Training Group interaction indicated that greater than expected improvement was shown by individually trained males, while females benefited more from group than individual training. Mean posttest scores of group trained males and individually trained males were 22.40 and 23.14, respectively. The girls' scores were generally lower, but those who were trained in groups attained slightly higher scores than those individually trained (20.13 and 18.89, respectively).

Means in the Race X Training Group interaction showed that there was a discrepancy between posttest scores of blacks and whites who were group trained (19.9 and 22.9, respectively) but no difference in posttest scores of individually trained blacks and whites (21.0 and 21.2, respectively). Whites in the control group attained a mean posttest score of 19.7 while the mean posttest score of black control subjects was only 17.8. Since both black and white control subjects had the same mean pretest score (17.7), white children gained more than black children merely from the experience of re-taking the test.

To examine the SES X Training Group interaction, SES was dichotomized into low SES versus others because of the markedly skewed distribution of SES level. The low SES group comprised children whose fathers were welfare recipients and unskilled workers (Categories 0 to 1 on the Turner Scale, 1964); the other SES group, which shall be referred to as "middle" SES, comprised the remainder of the sample which ranged from semi-skilled to professional workers (Categories 2 through 8). Posttest means of group and individually trained middle SES subjects and individually trained low SES subjects were similar (21.0, 21.5, and 21.2, respectively); the fact that the posttest mean of group trained low SES subjects was slightly lower (20.0) contributed to the variance accounted for by the SES X Training Group interaction ($\underline{r}^2 = .04$).

Means in the SES X Race interaction indicated a difference between posttest scores of black and white children at the lowest socioeconomic level (means = 18.3 for blacks and 21.3 for whites). At higher SES levels, means for blacks and whites were identical (20.4 and 20.3, respectively). The negative sign of the partial $\underline{r}$ for the SES X Race interaction is most likely due to the greater than expected improvement of low SES white children.

Discussion

The study has demonstrated that a group training procedure may be used with the Raven LP measure, and that, for students as a whole, results with special class students are similar to those that are achieved when these children are individually trained.

It is interesting to compare the results of this study with those of two previous investigations. Corman and Budoff (1973a) determined the effectiveness of group versus individual training with 202 intellectually normal children and found that children with low pretest scores benefited more from individual than from group training, when groups ranged from 12 to 15 students. These results suggested that a group training procedure might not be effective with special class students. Findings of the present study revealed that for the sample as a whole, the posttest means of the two trained groups were very similar, the mean gains of the two trained groups were more than double the mean gain of the control group, and the interaction between pretest score (R1) and Training Group was not significant. These results indicate that for special class students as a whole the two training procedures are equally effective when groups consist of five to eight students. Individual training was more effective than group training, however, for those students who were male, black, or in the lowest SES range.

Another investigation (Budoff & Corman, 1973) with 379 intellectually normal children, in the same school system from which the present sample was drawn, sought to determine the effects of race and social class on posttest scores of these children. The methodology used was parallel to that of the present study, except that half the normal students were trained in groups of 6 to 13 and the other half received no training. Significant main effects were found to be pretest score, age, SES, and training. Although 53% of this sample were black, in contrast to findings with the special class sample, the main effect for race and all interactions involving race were found to be not significant. Middle class students in the normal sample performed better on the posttest regardless of whether they were white or black. In this special class sample, on the other hand, the social class range was highly restricted and, without regard to race, social class was not a significant predictor of posttest performance. In this sample whites benefited more than blacks from both group training and practice in test taking. The discrepancy between posttest scores of blacks and whites was greater in the lowest socioeconomic range.

References

Budoff, M. Measuring learning potential: An alternative to the traditional intelligence test. Studies in Learning Potential, 1972, 3, No. 39.

Budoff, M., & Corman, L. The effectiveness of a group training procedure on the Raven Learning Potential Measure with children from diverse racial and socio-economic backgrounds. Studies in Learning Potential, 1973, 3, No. 58.

Budoff, M., Corman, L., & Litzinger, C. S. A Manual for the Raven Learning Potential Test. Studies in Learning Potential, 1973, 3, No. 59.

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven Learning Potential measure. Studies in Learning Potential, 1973, 3, No. 56. (a)

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven Progressive Matrices. Studies in Learning Potential, 1973, 3, No. 54. (b)

Turner, R. H. Social context of ambition: A study of high school seniors in Los Angeles. San Francisco: Chandler Publications, 1964.

Footnotes

[1] This research was supported by Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education, Grant NE-G-00-3-0016 from the National Institute of Education, and Grant R01 MH 18553 from the National Institute of Mental Health, all under the Department of Health, Education, and Welfare. Thanks and warm appreciation for their cooperation are due Dr. Bernice Kipfer, Director of Special Education, Mr. Arnold Berger, Director of Pupil Personnel Services, and the principals of the Syracuse, New York, Schools.

[2] 29 Ware Street, Cambridge, Massachusetts 02138

TABLE 1

Means and Standard Deviations of Three Groups on Raven Pretest and Posttest

Group	Pretest		Posttest		
	$\bar{X}$	SD	$\bar{X}$	SD	N
Group trained	17.929	5.861	21.389	7.268	56
Control	17.683	6.095	19.018	6.773	60
Individually trained	17.948	5.698	21.089	6.286	58

TABLE 2

Results of Multiple Regression on Posttest Scores

Variable	Beta	T-test	df
R1	.702	13.80***	157
Age	.137	2.74**	157
Sex	-.138	-3.06**	157
SES	-.045	-0.98	157
Race	.104	2.28*	157
Training condition (trained vs. nontrained)	-.148	-3.33**	157
Training group (group, control, or individual)	-.047	-1.04	157

$r^2 = .696$

$F = 51.34$, $df = 7/157$***

*$p < .05$

**$p < .01$

***$p < .001$

STUDIES IN LEARNING POTENTIAL

IQ AND LEARNING POTENTIAL MEASUREMENTS OF GENERAL INTELLIGENCE: A COMPARISON OF RELATIONSHIPS

By

Louise Corman and Milton Budoff

Volume 3, Number 55

1973

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

IQ AND LEARNING POTENTIAL MEASUREMENTS OF GENERAL INTELLIGENCE: A COMPARISON OF RELATIONSHIPS

Abstract

Demographic and psychometric factors related to pre- and posttraining LP scores were determined by multiple regression analyses. With all variables partialled out, pretraining scores were predicted by verbal and nonverbal IQs and demographic factors reflective of higher SES. Race, social class variables, and verbal IQ were not related to posttraining scores.

IQ AND LEARNING POTENTIAL MEASUREMENTS OF GENERAL INTELLIGENCE: A COMPARISON OF RELATIONSHIPS

Louise Corman and Milton Budoff
Research Institute for Educational Problems

Learning potential assessment represents a response to the growing dissatisfaction with the validity of traditional IQ scores for poor, non-middle-class, and frequently, non-white populations. IQ tests (group and individual) primarily assess the extent to which a child has spontaneously acquired knowledge and school-related proficiencies from his natural environment. The IQ test is based on the assumption that children have equal opportunities of access to school-preparatory experiences, and, thus, that differences in acquisition of knowledge, information, and skills reflect differences in (inborn) intellectual abilities. Unfortunately, poor and/or non-white children often lack access to appropriate school-preparatory experiences in early childhood, and perform poorly on IQ tests administered during the school years, indicating a low probability for success in academic school programs. If the IQ scores were viewed merely as predictors of scholastic success, these children's inferior IQs should indicate the need for altered school programs to forestall a predicted failure. Rather, the low IQ scores are usually interpreted to indicate lower (inborn) intelligence, and result in the misclassification of substantial proportions of non-middle-class children. It is underestimating their potential capabilities since their prior experiences have not prepared them for working with and solving the types of problems presented.

IQ testing is a particularly discriminatory practice when used as the primary criterion for classifying low income and/or non-white children as mentally retarded and segregating them into special classes, or into any other low educational track that may restrict the child's future educational opportunities. These children are grossly over-represented as educable retarded. More than 85% of all children in these classes are poor and/or non-white; the middle class children invariably have organic brain complications or severe emotional disturbance or both. Clearly a new approach to the testing of mental ability is required which would minimize the cultural biases of the tests, and offer insights into the potential abilities which could be incorporated into educational programs for these children.

The Research Institute for Educational Problems has been studying learning potential assessment as an alternative procedure for estimating general ability to reason. The learning potential (LP) assessment approach is based on a conceptualization of intelligence which stresses trainability, or the ability to profit from learning experiences. It assumes that the prior and continuing experiences of poor and/or non-white children do not allow them to spontaneously acquire school-relevant skills in their natural environment. The learning potential paradigm provides subjects with training experiences directly relevant to the reasoning task presented, so they can apply, in microcosm, their problem-solving ability and show whether they can improve their performance on the task. The improved performance indicates problem-solving capability not evident when the equalizing experiences of training are provided as part of the test administration.

Learning potential assessment replaces the traditional product-oriented test with a three-stage procedure which includes a pretest, one or more training sessions, and a posttest. Investigations of learning potential have usually employed the Kohs Block Designs or the Raven Progressive Matrices as pretest and posttest.

This paper summarizes data comparing the patterns of correlations of psychometric, social, and demographic variables with IQ and LP scores derived from the Kohs and Raven procedures, respectively, with large samples of educable mentally retarded (EMR) special class and institutionalized students. The purpose of these studies was to determine demographic and psychometric factors that accounted for significant portions of variance in scores on the Kohs and Raven problems before and after training, and to compare these factors with those related to scores on the Stanford-Binet IQ Test. The results of these two investigations will be presented separately.

Factors Related to Improved Performance after Training on the Kohs LP Measure

The sample for this study consisted of 627 EMRs from nine cities and towns in Massachusetts. Seventy-five percent (N = 471) were students in segregated special classes in public schools. Most of the remainder (N = 134) were residents of state institutions for the retarded; 22 were participants in a community workshop. The subjects ranged in age from eight to forty years, with a mean of 14.55 years (SD = 2.75) at the time of initial testing. Fifty-nine percent were males, 38% were white, and 79% had fathers who were manual laborers or menial service workers. Stanford-Binet IQ scores were obtained for 535 subjects; those scores ranged from 65 to 98, with a mean of 68.81 and standard deviation of 10.26.

Data on the following variables were collected from school or institutional records: place and date of birth, father's occupation, race, family size and degree of intactness, number of diseases, age at entry into special class, and WISC and Stanford-Binet IQs. Raven Progressive Matrices were group administered in the usual one session format by project staff. The Kohs learning potential measure was administered individually to each subject. The sixteen block designs were administered three times: prior to, and one day, and again one month following training. (See Budoff (1969) for details of administration.)

Stepwise multiple regression analyses were performed against four main dependent variables: Stanford-Binet IQ, pretraining score (K1), immediate posttraining score (K2) corrected by pretraining score, and delayed posttraining score (K3) corrected by the two prior Kohs scores. Major background predictors were evaluated both for their simple relationships to these four dependent variables and for their unique contributions to the variance of the dependent variables. Simple relationships between predictors and criteria were described by zero-order correlation coefficients. In order to test the unique contributions of predictors, the set of independent variables in question (e.g., all of the age-related variables) were forced into the regression equation after all of the other variables had been entered. From the remaining partials, one can infer each variable's unique contribution in predicting the dependent variable.

Sets of independent variables in the equations included chronological age, race (Black and Caucasian), social class, sex, birthplace, family size (number of children) and intactness, number of diseases, and scores on various psychometric measures. Turner's classification (1964) was used as the measure of social class.

Partial correlation coefficients indicated that the Stanford-Binet score was significantly related to non-institutionalization, family intactness, and WISC Verbal IQ, whereas partial r's showed that pretraining Kohs (K1) scores were not significantly related to any of these factors.

With all variables held constant, factors related to immediate effects of training on the Kohs were male sex; family size; birthplace outside Northeast U.S., or in foreign country; and high scores on the Stanford-Binet, Raven Progressive Matrices, and WISC Performance IQ tests. The only factors significantly related to delayed posttest scores obtained one month following training on the Kohs (K3 with all variables partialled out), were WISC Performance IQ and scores on the Raven Progressive Matrices. WISC Verbal IQ was not found to be uniquely related to any of the three Kohs scores. In several equations where race and social class were used as independent variables, these two factors were not found to be significantly related to Kohs posttraining scores.

Factors Related to Improved Performance on the Raven LP Measure

The sample for this study consisted of 403 EMR children who attended segregated special classes in Massachusetts public schools, including 130 children who had participated in the study previously described. They ranged in age from about seven to fifteen, with a mean age of 11 years (*SD* = 2). Sixty-two percent were male and 20% were Black. Three fourths of the subjects were from working class families. The mean Stanford-Binet IQ score was 70.95, with a standard deviation of 7.77. Raven Progressive Matrices (Sets A, A_B, B (1956) and Sets C, D, E (1958)) were group-administered by project staff before and after subjects received training. Training was done individually in two sessions. (See Budoff (1972) for details of administration.) The mean time interval between pre- and posttest was 30 days (*SD* = 29).

Insert Table 1 about here

Table 1

Partial Correlation Coefficients of Variables Predicting Stanford-Binet IQ, K1, K2, and K3

Variable Set	Variables Included	r_{SBIQ} zero order	r_{SBIQ} all variables partialled out	r_{K1} zero order	r_{K1} all variables partialled out	r_{K2} K1 partialled out	r_{K2} all variables partialled out	r_{K3} K1 and K2 partialled out	r_{K3} all variable partiall out
Age	Age at K1	-.402**	-.205**	.058	.248**	-.160**	.039	-.033	-.075
	Age first in special class	-.150**	-.091*	.014	.110*	-.170**	-.046	.019	-.031
	Years in special class	-.154**	-.010	-.029	-.003	.004	.031	-.061	-.011
Sex (1 = female, 2 = male)		.133**	-.027	.163**	.183**	.093*	.091*	.015	.050
Race (1 = white, 2 = black)		.086	.006	.026	.019	.045	not tested	-.055	not tested
Social class (1 = low, 9 = high)		-.009	.009	-.078	-.082*	.044	not tested	.058	not tested
Family size		-.024	-.030	.041	.033	.182**	.170**	-.025	-.002
Family intactness	Institutionalized (1 = yes, 2 = no)	.499**	.207**	.099*	-.009	.104*	-.074	.083	.075
	Father in home (1 = yes, 2 = no)	-.033	.072	-.001	033	-.078	-.028	-.073	-.071
	Mother in home (1 = yes, 2 = no)	-.057	-.007	-.020	.017	-.078	-.022	-.038	-.038

Table 1 (continued)

		r_{SBIQ}		r_{K1}		r_{K2}		r_{K3}	
Variable Set	Variables Included	zero order	all variables partialled out	zero order	all variables partialled out	K1 partialled out	all variables partialled out	K1 and K2 partialled out	all variables partialled out
	Intactness rating (1 = low, 9 = high)	.449**	.161**	.080*	.016	.129**	-.027	.116*	.076
Birthplace	Born in Northeast (1 = no, 2 = yes)	.236**	.003	.021	.003	.086*	-.084*	.007	.096
	Born in other U.S. (1 = no, 2 = yes)	.032	-.035	-.021	-.014	.000	-.023	-.081	-.040
	Foreign born (1 = no, 2 = yes)	.105*	.032	.024	.011	.169**	.148**	.070	-.093
Number of diseases		-.033	-.016	-.049	-.004	-.114*	-.053	.056	.005
Verbal IQ	Stanford-Binet IQ	---	---	.309**	.179**	.225**	.100*	.063	-.014
	WISC verbal IQ	.436**	.287**	.147**	-.076	.059	-.029	.059	-.010
Performance IQ	Raven score	.260**	.136**	.351**	.304**	.221**	.229**	.228**	.201**
	WISC performance IQ	.380**	.086*	.379**	.336**	.186**	.141**	.173**	.152**
N		535	535	627	627	627	627	409	409

*p <.05

**p <.01

The following factors were examined with respect to their unique contribution to pre- and posttraining scores on the Raven: social class, race, sex, age, number of diseases, family size, family intactness, birthplace, score on a group IQ test, WISC Verbal IQ, WISC Performance IQ, and length of time interval between pretest and posttest. Statistical methodology was parallel to that used in the study with the Kohs LP measure. Six multiple regression equations were performed with the following Raven scores as dependent measures: pretraining scores on Sets A, A_B, B; pretraining scores on Sets C, D, E; pretraining scores on the total test; and three posttraining scores for these sets corrected by the corresponding pretraining scores.

Results indicated that variables uniquely related to a high pretest score on the total Raven test were: male sex, age, few older siblings, father absent from the home, and a high Stanford-Binet and WISC Performance IQ. Greater than expected improvement following training appropriate to these problems on Sets A, A_B, B was uniquely related only to being male, having a high WISC Performance IQ, and long interval between test sessions. Comparison between factors uniquely related to performance before and after training indicated that effects due to age, number of older siblings, father's absence, and Stanford-Binet IQ, which were related to pretest performance, were not uniquely related to improved scores on the total Raven test after training.

Insert Table 2 about here

Table 2

Partial Correlation Coefficients of Variables Predicting AABB, CDE, and Total Raven Pretest and Posttest

Variable set	Variables included	r_{AABB1}		r_{AABB2}		r_{CDE1}		r_{CDE2}		r_{R1}		r_{R2}	
		Zero order	Partial r^a	Partial r^b	Partial r^a	Zero order	Partial r^a	Partial r^b	Partial r^a	Zero order	Partial r^a	Partial r^b	Partial r^a
Social Class (1 = low; 9 = high)		-.060	-.006	-.107*	-.106*	.044	.000	.066	.083	.029	-.003	.031	.051
Race (1 = white;2 = black)		.029	.027	.069	.055	-.116	-.081	.018	.000	-.113	-.087	.043	.040
Sex (1 = male; 2 = female)		-.239**	-.123*	-.156**	-.111*	-.220**	-.107	-.078	-.006	-.287**	-.153*	-.194**	-.126*
Age		.500**	.458**	.077	.087	.414**	.392**	.226**	.208**	.478**	.494**	.114	.101
Diseases		-.175**	-.021	-.131**	-.040	-.002	-.002	-.108	-.013	-.065	.004	-.085	.003
Family Size	Older Siblings	-.077	-.093	-.021	-.038	-.077	-.091	-.081	-.141*	-.109	-.143*	.084	-.119
	Younger Siblings	.183**	.125*	.078	.104*	.163**	.086	.078	.056	.183**	.101	-.058	.113
Family Disintegration	Father Absent (1=no;2=yes)	.127*	.109*	.000	.035	-.057	.108	.054	.049	.143*	.144*	.036	.072
	Mother Absent (1=no;2=yes)	.020	.018	.139**	.152**	.113	-.063	.027	.011	-.025	-.023	.126*	.120
	Disintegration Rating (1 = low; 9 = high)	.026	.040	.022	.039	-.041	.036	-.064	-.035	-.025	.062	-.038	.020

Table 2 (continued)

Variable set	Variables included	r_{AABR1}		r_{AABB2}		r_{CDE1}		r_{CDE2}		r_{R1}		r_{R2}	
		Zero order	Partial r^a	Partial r^b	Partial r^a	Zero order	Partial r^a	Partial r^b	Partial r^a	Zeor order	Partial r^a	Partial r^b	Partial r^a
Birthplace	Born in West or South U.S. (1 = no; 2 = yes)	-.036	-.067	.007	-.002	-.085	-.031	.000	.036	-.091	-.037	.018	.043
	Foreign Born (1=no; 2=yes)	.099*	-.005	.087	.097	.095	-.017	.030	.060	.108	-.004	.058	.099
Group IQ		.090	.010	-.003	-.003	.035	.054	.008	.035	.073	.121	-.079	-.066
Verbal IQ	Stanford-Binet IQ	.207**	.152**	.065	.016	.197**	.183**	.042	.056	.254**	.251**	.039	.022
	WISC Verbal IQ	.169**	.040	.087	.030	.017	-.089	.003	.000	.042	-.066	.059	.014
Performance IQ	WISC Performance IQ	.354**	.261**	.207**	.163**	.309**	.290**	.116	.113	.351**	.327**	.176**	.142*
Days from R1 to R2		---	---	.054	.072	---	---	.189**	.155*	---	---	.185**	.149*
N		403	403	403	403	266	266	259	259	266	266	259	259

[a] Partial r obtained with all other variables partialled out.

[b] Partial r obtained with only corresponding pretest score partialled out.

*$p < .05$

**$p < .01$

Regression analyses with posttraining scores on Sets C, D, E as the dependent variable indicated that older children and children with fewer older siblings improved the most on these more difficult items. For the total sample, social class and being white were not positively related to any pre- or posttest scores on the Raven test.

With all other variables held constant, WISC Performance IQ was found to predict significantly all pre- and posttest scores, with the exception of posttraining scores on Sets C, D, E. By comparison, WISC Verbal IQ was not significantly related to any pre- or posttest scores, when all other variables were partialled out.

Summary

The IQ test, which measures the child's spontaneously acquired skills which predict academic outcomes, misclassifies disproportionate numbers of poor and/or nonwhite populations as mentally retarded. Since prior experiences may vary, learning potential assessment, which uses a (multi-session) test-train-test paradigm, defines intelligence as ability to profit from suitable experience. Low IQ students (60 to 80) demonstrate marked heterogeneity of ability on a training-based assessment procedure which teaches principles relevant to the reasoning task (Kohs Block Designs or Raven Progressive Matrices). Data were presented comparing the pattern of correlations of psychometric, social, and demographic variables with IQ and LP scores, respectively. Variables commonly

associated with socioeconomic status were correlated with IQ and pretest administration of the LP measure. Following training, LP scores were related to performance IQ but unrelated to race, social class, or verbal IQs.

References

BUDOFF, M. (1969) Learning Potential: A supplementary procedure for assessing the ability to reason. *Seminars in Psychiatry*, *1*, 278-290.

BUDOFF, M. (1972) Providing special education without special classes. *Journal of School Psychology*, *10*, 199-205.

RAVEN, J. C. (1956) *Coloured progressive matrices: Sets A, A_B, B*. London: H. K. Lewis & Co., Ltd.

RAVEN, J. C. (1958) *Standard progressive matrices, Sets A, B, C, D, and E*. London: H. K. Lewis & Co., Ltd.

TURNER, R. H. (1964) *Social context of ambition: A study of high school seniors in Los Angeles*. San Francisco: Chandler Publications.

Footnotes

[1] This work was made possible by Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education, and Grant NE-G-00-3-0016 from the National Institute of Education, and Grant R01 MH 18553 from the National Institute of Mental Health. The paper was prepared and presented at the Third International Congress of the International Association for the Scientific Study of Mental Deficiency, The Hague, The Netherlands, 4th - 12th September, 1973.

[2] The Research Institute is located at 29 Ware Street, Cambridge, Massachusetts 02138, U.S.A.

STUDIES IN LEARNING POTENTIAL

FACTORIAL COMPOSITION OF RAVEN'S PROGRESSIVE MATRICES IN THE CONTEXT OF LEARNING POTENTIAL ASSESSMENT

Louise Corman and Milton Budoff

Volume 3, Number 62

1973/1975

RESEARCH INSTITUTE FOR EDUCATIONAL PROBLEMS

29 WARE STREET
CAMBRIDGE, MASSACHUSETTS 02138
617-868-0360

FACTORIAL COMPOSITION OF RAVEN'S PROGRESSIVE MATRICES IN THE CONTEXT OF LEARNING POTENTIAL ASSESSMENT

Abstract

The purpose of this study was to examine the stability of the factor structure of the Raven Progressive Matrices when problem-relevant training was interpolated between two administrations of the test. Raven's Coloured Progressive Matrices (Sets A, A_B, B) were administered twice to normal and Educable Mentally Retarded (EMR) subjects in two samples which differed in racial composition. Approximately half the subjects received problem-relevant training between the two test sessions. Factor analysis indicated comparability of factor structures of normal and EMR subjects before and after retest with or without training. Varimax rotations of all tests performed separately with two samples (*N* = 1238 and 1089 tests) revealed factor structures which were nearly identical. This study provides evidence of the factorial invariance of the Raven test.

FACTORIAL COMPOSITION OF RAVEN'S PROGRESSIVE MATRICES IN THE CONTEXT OF LEARNING POTENTIAL ASSESSMENT

Learning potential (LP) assessment represents an alternative procedure to traditional IQ tests for estimating general ability to reason. Originally conceived as a more equitable means of discriminating ability to reason among severely low achieving school children who had been classified as educable mentally retarded (EMR), the learning potential assessment has been used successfully with black, white, and Spanish-speaking higher IQ children who are in high educational risk. This assessment approach is based on a conceptualization of intelligence which stresses trainability, or the ability to profit from learning experiences. It assumes that the prior and continuing cognitive and motivational experiences of poor and/or nonwhite or non-English-speaking children inhibit them from spontaneously acquiring or applying school-relevant skills in their natural environment. The learning potential paradigm provides subjects with training experiences directly relevant to the reasoning task presented, so they can apply, in microcosm, their problem-solving ability and demonstrate whether they can profit from the training experience by improving their performance on the task. Learning potential assessment replaces the traditional product-oriented test with a three-stage procedure which includes a pretest, a training session, and a posttest.

Budoff (1969, 1972) has argued that providing training on the task-relevant skills to low socioeconomic status children will diminsh the competence gap between them and middle class students by equalizing the child's experience in these types of problem-solving tasks with those of middle class children. We hypothesize that on reasoning types of problems (Level 2) the lower scores commonly associated with differences in socioeconomic status (Jensen, 1969) may be due to the discontinuity between the life experiences of low-SES children and the school-relevant skills found in traditional intelligence measures. These latter cognitive and perceptual skills are spontaneously acquired and applied by middle class children, and these children do obtain higher scores on the tests and do well in school.

Budoff and Corman (1973) reported that 40% of students from lower income homes attained scores on Raven's Progressive Matrices which were as high or higher than the pretrained mean of middle class students. Babad and Budoff (in press) and Corman and Budoff (1973a) reported similar findings with a different nonverbal reasoning task (a series test) which was administered in the LP test-train-retest paradigm. Babad and Budoff (1974) included samples of students from special classes for the educable mentally retarded (IQ < 80), and those with IQs in the 80 to 100, and 100 plus ranges. Sixteen percent of the educable mentally retarded sample fell at or above the pretest mean of the dull normal sample (mean IQ difference was 17 points) and 36% exceeded this pretest mean after training. With the bright normal group as criterion for the EMR subjects (there was a 45 point difference between the mean IQs of the two groups), none of the EMR children reached the bright normal mean in the initial test; 13% exceeded it following training. Eleven and 35% of the dull normal group attained the pretraining mean of the bright normal group prior to and following training. Corman and Budoff (1973a) reported that 51.3% of their low income normal sample exceeded the middle class pretest mean after training was provided. What is notable in this study is that the low-SES sample was largely black (approximately 80%) and the contrast middle SES-students lived in affluent white communities in western Connecticut.

Raven's Coloured Progressive Matrices (1956) has been widely used as a measure of nonverbal reasoning ability and was employed as the measure of learning potential in this study. This test consists of 36 six-choice problems comprising Sets A, AB, and B with 12 problems in each set. Raven (1965) has described seven types of problems represented by the items in these sets as follows: (1) problems requiring simple continuous pattern completion (A1 to A8), (2) pattern completion with change in one direction (A9 to A10), (3) pattern completion with change in two directions (A11 to A12), (4) discrete pattern completion (AB1 to AB3, B1 to B2), (5) apprehension of three figures as a whole (AB4 to AB12, B3 to B5), (6) concrete reasoning by analogy (B6 to B9), and (7) abstract reasoning by analogy (B10 to B12).

Corman and Budoff (1974) examined the factor structures of normal and EMR children on a single administration of the Raven test and identified four factors: I Continuity and Reconstruction of Simple and Complex Structures, II Discrete Pattern Completions, III Reasoning by Analogy, and IV Simple Continuous Pattern Completion. Factors II, III, and IV closely resembled certain item categories postulated by Raven (1965). Factor II was very similar to Raven's category 4; Factor III included most of the items in categories 6 and 7; and Factor IV was similar to Raven's category 1. This factor structure was obtained with both normal and EMR children and was shown to be replicable with two independent samples.

In the present study, factor structures of normal and EMR children on tests taken after learning potential traing or after retesting with no intervention were initially compared with factor structures obtained with the same subjects on the pretest. This preliminary analysis sought to determine whether training or practice markedly changed the factor structure obtained on the initial test administration. If factor structures after training or practice were similar to the pretest factor structure perviously identified (Corman & Budoff, 1974), the factorial composition of all pre- and post-tests administered in LP assessment could be examined as a whole.

Factor structures were obtained with two samples which differed in racial composition, so that the replicability of the factor structures could be examined. These two samples of pretest factor structures permitted comparison between a predominantly white sample and a sample in which approximately 50% of the students were black. Investigations with the two samples will be referred to as Study 1 and Study 2.

Method

Subjects

The sample in Study 1 consisted of 243 normal and 399 EMR subjects. Normal subjects were drawn from low income areas of a large urban community in Massachusetts. They ranged in age from seven to twelve and were in the second through sixth grade. Approximately 60% were male and about a fourth were black. EMR subjects ranged in age from seven to fifteen, with a mean age of eleven years (SD = 2). Sixty percent were male and about 25% were black. Three-fourths of the EMR subjects were from working class families. The mean Stanford-Binet IQ of EMR subjects for whom this data was available (N = 242) was 71.0, SD = 7.8. All EMR subjects attended special classes in public schools.

The sample in Study 2 consisted of 379 normal and 174 special class students who attended schools in a large urban community in upstate New York. Normal subjects ranged in age from six to eleven and were evenly distributed among the first through fifth grades. Fifty-three percent of these subjects were black and 28% of the normal students' fathers were employed in business or professional occupations; the majority were blue collar or clerical workers. EMR subjects ranged in age from 5-1/2 to 14 years, with a mean age of 10 (SD = 2). All but 5% were from low socioeconomic backgrounds (welfare, blue collar, or clerical) as indicated by father's occupation, and 43% were black. The mean IQ of the EMR subjects on the WISC or Stanford-Binet was 68.4, with a standard deviation of 9.0. Fifty-five percent of both normal and EMR subjects were male.

The samples for Studies 1 and 2 were similar in most respects other than racial composition; approximately half of the subjects in Study 2 were black, in comparison to one-quarter of the Study 1 sample.

Procedure

In Study 1, intact classrooms of normal children were randomly assigned to either a nontrained group or to a group in which each member was individually trained on problems similar to those on the Raven test. Three hundred fifty-three EMR subjects received individualized LP training, and 46 EMR children were selected at a later date to constitute a nontrained group. Pretests were group administered to the total sample of 243 normal and 399 EMR subjects and posttests were group administered to 147 trained normal, 84 nontrained normal, 319 trained EMR, and 46 nontrained EMR subjects. (The remaining 12 normal and 34 EMR subjects were absent during posttesting.) Thus, the total number of tests administered was 1238. The interval between pretest and posttest varied from two to 180 days but averaged approximately 30 days, with a standard deviation of about 29 days.

In Study 2, both normal and EMR subjects were assigned to trained and nontrained groups with a randomized block procedure at the outset of the study. On the basis of pretest scores, normal subjects were randomly assigned to a nontrained group or to a group which received LP training in groups of six to thirteen students. EMR subjects were randomly assigned to one of three groups on the basis of their pretest scores: a nontrained group, a group which received individual training, and a group which received training in subgroups of five to eight students. Since analysis of variance revealed no significant difference between posttest scores of EMR subjects who were trained individually or in a group, subjects who received both types of training were combined to constitute the trained group of EMR children.

A total of 1089 tests were administered to children in classroom-size groups. Pretests were given to the total sample of 379 normal and 174 EMR children, and posttests were administered to 172 trained and 199 nontrained normal subjects and to 110 trained and 55 nontrained EMR subjects. All posttests were administered within two weeks after pretests had been given. The training procedure is described elsewhere (Budoff, Corman, & Litzinger, 1973)

Results

Preliminary Analyses

1. Test-retest coefficients. To examine the stability of the distribution of total scores over time, test-retest correlations were obtained separately for EMR and normal children in the trained and nontrained groups in Studies 1 and 2. Correlations for nontrained subjects represent test-retest reliability coefficients. The test-retest correlations presented in Table 1 indicated the stability of the distribution of total scores over time for both samples of trained normal subjects in Study 1 was lower than that obtained in Study 2. These coefficients indicated that training did not greatly affect the distribution of scores from pretest to posttest, particularly for the EMR samples. The test-retest reliability of the Raven test was quite high with both samples of normal children (.82 and .84). The test-retest coefficient for EMR children was high in Study 2 (.87) but only .51 in Study 1. The higher reliability obtained in Study 2 may have resulted from the shorter time interval between pre- and posttest in that study. The fact that coefficients for trained subjects were generally somewhat lower than those for nontrained subjects indicates that training affected the distribution of total scores to a greater extent than did practice.

Insert Table 1 about here

----

2. Factor analysis. Statistical analysis was identical in Studies 1 and 2. Pretest and posttest responses of all subjects to the 36 six-choice items of Sets A, AB, B were dichotomized into correct and incorrect responses (1 = incorrect, 2 = correct). The 36 X 36 item intercorrelation matrix was factored by the principal components method with 1.00 in the diagonal. On the basis of previous factor analytic findings with the Raven test (Corman & Budoff, 1973b), four factors were specified for the subsequent normal varimax rotations. This procedure was carried out separately with six sets of tests: pretests of normal children, pretests of EMR children, posttests of trained normal subjects, posttests of nontrained normal subjects, posttests of trained EMR subjects, and posttest of nontrained EMR subjects.

The factor structures were then compared for the six sets of tests in Studies 1 and 2 (see Appendix A). Four factor solutions on the six separate sets of tests met the criteria of simple structure and satisfied the criteria recommended by Kaiser (latent roots > 1.00) and by Cliff and Hamburger (1967)

TABLE 1

Test-Retest Correlation Coefficients for EMR and Normal Subjects in Studies 1 and 2

	Study 1	N	Study 2	N
EMR				
Trained	.751	319	.769	110
Nontrained	.514	46	.871	55
Normal				
Trained	.605	147	.749	173
Nontrained	.817	84	.839	201

(significant breaks in latent roots). The four factors accounted for 38 to 51% of the total variance in the six analyses with the samples in Study 1 and 40 to 53% of the samples in Study 2. The four factors which appeared in the pretest analysis of normal children in Study 1 were considered as referent factors in the other analyses.

While these referent factors accounted for different amounts of variance in the six analyses, certain items had high loadings on the same reference factor in all analyses with samples in both studies (Table A3). Factors I and III contained many of the same items in all six analyses. Factor IV contained fewer items with high loadings when both samples are considered, but nevertheless represented a fair degree of consistency across the six sets of tests. Factor II was least consistent. Items comprising this factor appeared in analyses of pretests taken by normal and EMR children, as well as posttests of untrained normal and trained EMR children; however, Factor II was not present in analyses of posttests of trained normal and untrained EMR subjects. Difficulty levels of items in Factor II indicated that these were easy items for both normal and EMR children and resulted in little variability. These comparisons can be examined in Table A1 - A3 in the appendix.

Training and practice did not appear to markedly change the Raven factor structure. The factor structures on the six sets of tests were considered comparable on the whole, and all tests were pooled into one sample for Study 1 (1238 tests) and a second sample for Study 2 (1089 tests). A principal components analysis with 1.00 in the diagonal, followed by a varimax rotation were performed with both test samples.

3. Definition of the factors. Table 2 presents the items with high loadings on the four factors obtained in Study 1 and 2. The magnitude of the first seven latent roots obtained in the principal components analysis with the sample in Study 1 (9.23, 3.31, 1.80, 1.34, 1.06, 1.00, and .94, respectively) indicated that the roots of the first four factors were the most significant by the criteria perviously described. For this reason, the four factor solution was considered most appropriate for this sample. These first four factors accounted for 43.6% of the total variance. The first four latent roots obtained in the principal components analysis with the 1089 tests in Study 2 were comparable in magnitude to those contained in Study 1 (8.70, 3.29, 1.74, and 1.37, respectively). Consequently, the percent of variance accounted for by all four factors and the percent accounted for by each factor were similar to the corresponding figures in Study 1. The four factors obtained in Study 2 accounted for 42.0% of the variance compared to 43.6% in Study 1.

Insert Table 2 about here

The four factors corresponded to readily identifiable types of problems in the Raven test. Factor loadings with the sample in Study 1 indicated that Factor I consisted of Items A7 to A12, AB4 to AB11, B3 to B5, and B7, all of which are at a moderate difficulty level. Items A7 to A12 involve the ability to discern continuity in a complex whole; Items AB4 to AB7 and B3 to B5 represent reconstruction of simple discontinuous structures; Items AB8 to AB11 represent reconstruction of structures based on symmetry; and Item B7 involves reconstruction of complex structures. Factor I can be called Continuity and Reconstruction of Simple and Complex Structures.

Factor II consisted of Items AB1 to AB3 and B1 to B3, which are all items requiring simple completions. These items involve the ability to discern similarity and continuity of a simple geometric background. With the exception of Item B3, which loaded more highly on Factor I, Factor II is identical to Raven's Category 4, Discrete Pattern Completions.

Factor III comprised Items AB12, B6, and B8 to B12, all of which involve Reasoning by Analogy. Most of these items are included in Raven's Categories 6 and 7 (Concrete and Abstract Reasoning by Analogy).

Factor IV comprised Items A1 to A6. These are the easiest items on the test and consist of linoleum block completions. The six items in Factor IV are all included in Raven's Category 1, along with Items A7 and A8. Raven called these items Simple Continuous Pattern Completions.

When factor loadings in Study 2 are compared with those in Study 1, it becomes clear that the factor structure obtained with these two samples was nearly identical. Item B6 was the only item with a high loading on a particular factor in Study 2 but not Study 1; the loading of this item was .34 on Factor I in Study 1 and .41 in Study 2. Burt's coefficient of congruence (Harman, 1967) was used to determine the degree of relationship between loadings on each factor with the two samples. Coefficients for all four factors were very high (.97, .96, .98, and .94 for Factors I through IV, respectively), indicating comparability of the factor structures for these samples.

Discussion

Factor analysis of the Raven Learning Potential measure revealed a pattern of test items which could be logically as

TABLE 2

Loadings above .40[a] on Varimax Rotation of Raven Items in Studies 1 and 2[b]

Item	Factor I		Factor II		Factor III		Factor IV	
	1	2	1	2	1	2	1	2
A1							-.57	.40
A2							-.71	.49
A3							-.73	.72
A4							-.64	.72
A5							-.53	.57
A6							-.46	.55
A7	.56	.64						
A8	.44	(.36)						
A9	.55	.52						
A10	.58	.57						
A11	.54	(.39)						
A12	.42	(.38)						
AB1			.72	.72				
AB2			.66	.73				
AB3			.68	.73				
AB4	.54	.61						
AB5	.64	.64						
AB6	.68	.67						
AB7	.58	.62						
AB8	.53	.59						
AB9	.53	.57						
AB10	.48	.47						
AB11	.41	(.38)						

TABLE 2 (continued)

	Factor I		Factor II		Factor III		Factor IV	
Item	1	2	1	2	1	2	1	2
AB12					.60	.57		
B1			.65	.65				
B2			.63	.54				
B3	.46	.54	.40	.40				
B4	.59	.60						
B5	.61	.59						
B6		.41			.41	(.38)		
B7	.40	.45						
B8					.75	.74		
B9					.78	.72		
B10					.73	.70		
B11					.78	.79		
B12					.62	.60		
Rotated SS	5.788	5.976	3.425	3.147	3.891	3.593	2.575	2.398
% total variance	25.6	24.2	9.2	9.1	5.0	4.8	3.7	3.8
Cumulative %	25.6	24.2	34.8	33.3	39.8	38.2	43.6	42.0
Burt's coefficient	.97		.96		.98		.94	

[a] Loadings between .35 and .39 in Study 2 are in parentheses.

[b] N = 1238 tests in Study 1 and 1089 tests in Study 2.

well as statistically defined. The factor structure obtained with pretests (Corman & Budoff, 1974) did not differ markedly after training or after retesting with no intervention. Items grouped together in each factor corresponded to readily identifiable types of problems on this test. Four factors were obtained: I Continuity and Reconstruction of Simple and Complex Structures, II Discrete Pattern Completions, III Reasoning by Analogy, and IV Simple Continuous Pattern Completions. Factors II, III, and IV closely resembled types of items postulated by Raven (1965) which have been previously described.

Perhaps the most important finding of this investigation is that the factor structure of item responses on all pre- and posttests taken by EMR and normal subjects was nearly identical in Studies 1 and 2. Samples in the two studies differed primarily in racial composition; one-quarter of the subjects in Study 1 were black, while about half of the Study 2 sample was black. The independent replication of the factor structure with these two samples provides evidence of the factorial invariance of the Raven Progressive Matrices, even after problem-relevant training, with moderate variations of geographic conditions and racial composition.

References

Babad, E.Y., & Budoff, M. Sensitivity and validity of learning potential measurement in three levels of ability. *Journal of Educational Psychology*, in press.

Budoff, M. Learning potential: A supplementary procedure for assessing the ability to reason. *Seminars in Psychiatry*, 1969, *1*, 278-290.

Budoff, M. Measuring learning potential: An alternative to the traditional intelligence test. *Studies in Learning Potential*, 1972, *3*, No. 39.

Budoff, M., & Corman, L. The effectiveness of a group training procedure on the Raven learning potential measure with children from diverse racial and socioeconomic backgrounds. *Studies in Learning Potential*, 1973, *3*, No. 58.

Budoff, M., Corman, L., & Litzinger, C.S. A manual for the Raven learning potential test. *Studies in Learning Potential*, 1973, *3*, No. 59.

Cliff, N., & Hamburger, C.D. The study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin*, 1967, *68*, 430-445.

Corman, L., & Budoff, M. The series test as a measure of learning potential. *Studies in Learning Potential*, 1973, *3*, No. 47. (a)

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven Progressive Matrices. *Journal of Educational and Psychological Measurement*, 1974, *34*, 407-412.

Harman, H.H. *Modern factor analysis*. Chicago: The University of Chicago Press, 1967.

Jensen, A.R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, *39*, 1-123.

Raven, J.C. *Coloured progressive matrices: Sets A, AB, B.* London: H.K. Lewis & Co., 1956.

Raven, J.C. *Guide to using Raven Progressive Matrices*. London: H.K. Lewis & Co., 1965.

Footnotes

[1] This research was supported by grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education and grant NE-G-00-3-0016 from the National Institute of Education, both under the Department of Health, Education, and Welfare. The statistical consultation of Robert H. Harrison is gratefully acknowledged.

[2] Tables of items with high factor loadings on varimax rotations with the six sets of tests are available from the authors upon request.

APPENDIX A

Raven Items with Factor Loadings above .30 on Varimax Rotation of Six Sets of Tests in Studies 1 and 2

TABLE A1

Study 1: Raven Items with Loadings above .30 on Varimax Rotations for 6 Sets of Tests

	Normals			EMRs		
	Pretest	Posttest		Pretest	Posttest	
		Trained	Untrained		Trained	Untrained
Factor I	A7-A12 AB4-AB11 B3-B7	A1-A6 AB1-AB3 B1	A12 AB9-10, AB12 B5-B12	A7-A11 AB4-AB11 B3-B7	AB10, AB12 B6-B12	A2-4, A6-7 AB1-5, AB7, AB12 B1-B3
Referent factor	I	IV, II	III	I	III	IV, II
% variance	21.5	24.9	27.1	20.1	25.6	24.4
Factor II	AB1-AB3 B1-B2	A10 AB4-6, AB8 B2-B3, B5	A10-A11 AB3-AB8 B3-B4	AB12 B8-B12	A1-A6	AB8, AB11 B6-B12
Referent factor	II	I	I	III	IV	III
% variance	11.0	11.3	9.8	8.6	9.6	10.8
Factor III	AB12 B8-B12	A7 AB12 B4, B6, B8-12	AB2-AB3 B2	A2-A5	AB1-AB4 B1-B2	A5, A9-A11 AB6, AB9 B4-B5
Referant factor	III	III	II	IV	II	I
% variance	5.7	6.0	8.0	5.5	5.1	7.1
Factor IV	A1-A6	A8-9, A11-12 AB9-AB11	A5-A8, A11 AB11	A6 AB1-AB3 B1-B2	A7-A12 AB5-9, AB11 B3-B5	A1, A8, A12 AB10
Referent factor	IV			II	I	
% variance	4.3	4.7	5.8	4.3	4.0	5.9
% total variance	42.6	46.8	50.6	38.5	44.3	48.1
N	243	147	84	399	319	46

TABLE A2

Study 2: Raven Items with Loadings above .30 on Varimax Rotation

	Normals			EMRs		
	Pretest	Posttest		Pretest	Posttest	
		Trained	Untrained		Trained	Untrained
Factor I	AB12 B5-B12	A7, A9-A12 AB5-6, AB8-12 B5-B12	A7, A11 AB8-AB12 B5-B12	A7, A9-A11 AB4-AB10 B3-B6	A3, 7, A10-12 AB4, 6-7, 9-11 B5-B10	A5-A6, A9 AB1-AB4, AB11 B1-B2, B4, B9
Referent factor	III	III	III	I		II
% variance	23.0	25.5	23.6	20.1	23.4	25.2
Factor II	A4-A6, A8-A9 AB1-AB4 B1-B4	A2-A6, A8	A1, A3-6, A9 AB2 B1	A3-A6 AB1-AB3 B1-B3	A5, A8-10 AB1-3, AB7 B1-B3	A7-A8, A10, A12 AB4-AB10 B2-B7
Referent factor	II	IV	IV	II, IV	II	I
% variance	9.7	9.4	9.2	9.5	7.4	13.4
Factor III	A7-A12 AB4-AB11 B3-B7	A2, A7, A10 AB1-AB10 B5	A7-A10 AB4-9, AB11 B1-B7	AB12 B8-B12	A4, A7-9 AB8, AB12 B4, B8-12	A3-A5 AB12 B11-B12
Referent factor	I	I	I	III	III	IV
% variance	5.2	5.9	5.6	6.1	5.4	7.8
Factor IV	A1-A6	A7, A9-A10 B1-B5, B10	A2 AB1-AB3	A5-10, A12 AB4-5, AB11 B7	A3, A6 AB1, AB4-10 B3-B6, B10	A11-A12 AB6, AB8, AB12 B5-B6, B8-B12
Referent factor	IV	II	II		I	III
% variance	4.3	5.0	5.1	4.4	4.5	6.2
% total variance	42.2	45.8	43.4	40.1	40.7	52.6
N	399	172	199	174	110	55

TABLE A3

Raven Items with Factor Loadings above .30 on Varimax Rotation in Two Studies

	Normals			EMRs		
	Pretest	Posttest		Pretest	Posttest	
		Trained	Untrained		Trained	Untrained
Factor I	A7-A12 AB4-AB11 B3-B7	A10 AB4-AB6, AB8 B5	A10 AB4-AB8 B3-B4	A7, A9-A11 AB4-AB10 B3-B6	AB5-AB9 B3-B5	A10 AB6, AB9 B4-B5
Factor II	AB1-AB3 B1-B2	A9	AB2-AB3	A6 AB1-AB3 B1-B3	AB1-AB3 B1-B2	
Factor III	AB12 B8-B12	A7 AB12 B6, B8-B12	AB9-AB10, AB12 B5-B12	AB12 B8-B12	AB12 B8-B12	AB8 B6, B8-B12
Factor IV	A1-A6	A2-A6	A5	A4-A5	A3	A3-A4

riep **PRINTS**

FACTOR ANALYSIS OF ERROR TYPES ON RAVEN'S PROGRESSIVE MATRICES: A FAILURE TO OBTAIN LOGICAL STRUCTURE

Milton Budoff and Louise Corman

- RIEP-Print Number 79

June, 1975

RIEP-Prints are a continuation of the
series STUDIES IN LEARNING POTENTIAL

RESEARCH INSTITUTE FOR EDUCATIONAL PROBLEMS, INC.
29 WARE STREET ▪ CAMBRIDGE, MASSACHUSETTS 02138 ▪ TELEPHONE (617) 868-0360-2

© 1975

Factor Analysis of Error Types on Raven's Progressive Matrices: A Failure to Obtain Logical Structure[1]

Milton Budoff and Louise Corman

Widely used as a measure of reasoning in children, Raven's Coloured Progressive Matrices were designed to permit qualitative comparisons of erroneous responses which could indicate "erroneous" mental processes in the person taking the test (Raven, 1965). Sets A, AB, B, which comprise the Coloured Matrices, were developed for use with children up to approximately 11 years of age and children whose intellectual functioning has become impaired. Raven constructed the distractors of the 36 items in these sets to reflect 10 types of errors that could be made by the examinee.

This paper presents data collected from IQ-defined educable mentally retarded (EMR) and normal children under two conditions of experience to determine whether Raven's logically-defined factors group in interpretable empirically-defined factors. The use of children who vary in age, IQ, and experience on the task was intended to test the generalizeability of the separate solution(s) obtained.

Experience on the task was varied by interpolating training relevant to the item types between two administrations of the task to the average and low IQ children.

Method

Subjects

The sample consisted of 243 normal and 399 EMR subjects. Normal subjects were drawn from low income areas of a large urban community in Massachusetts. They ranged in age from seven to twelve and were in

Table 1

Description of Raven's (1965) 10 Error Types

Error type	Total number of errors possible on 36 items (A, AB, B)	Description
A	11	Difference: the piece has no figure of any kind on it.
B	38	Difference: the figure shown is quite irrelevant.
C	22	Inadequate Individuation: the figure is contaminated by irrelevancies or distortions.
D	5	Inadequate Individuation: it combines figures irrelevantly.
E	12	Inadequate Individuation: it is the whole or half of the pattern to be completed.
F	13	Repetition of the pattern: above and to the left of the space to be filled.
G	20	Repetition of the pattern: immediately above the space to be filled.
H	19	Repetition of the pattern: immediately to the left of the space to be filled.
I	14	Incomplete correlate: the figure is wrongly oriented.
J	26	Incomplete correlate: it is incomplete but correct as far as it goes.

the second through sixth grade. Approximately 60% were male and about a fourth were black. EMR subjects ranged in age from seven to fifteen, with a mean age of eleven years (SD =2). Sixty percent were male and about 25% were black. Three fourths of the EMR subjects were from working class families. The mean Stanford-Binet IQ of EMR subjects for whom this data was available (N = 242) was 71.0, SD = 7.8. All EMR subjects attended special classes in public schools.

Procedure

Intact classrooms of normal children were randomly assigned to either a non-trained group or to a group in which each member was individually trained on problems similar to those on the Raven test. Three hundred fifty three EMR subjects received individualized LP training, and 46 EMR children were selected at a later date to constitute a non-trained group. Pretests were group administered to the total sample of 243 normal and 399 EMR subjects and posttests were group administered to 147 trained normal, 84 non-trained normal, 319 trained EMR, and 46 non-trained EMR subjects. (The remaining 12 normal and 34 EMR subjects were absent during posttesting.) The interval between pretest and posttest varied from two to 180 days but averaged approximately 30 days, with a standard deviation of about 29 days. The training procedure is described in Budoff, Corman, and Litzinger (1975).

Analysis

The ten types of errors postulated by Raven are described in Table 1. The maximum possible frequency with which each error type can be made varies; there are six choices per item, one of which is the correct answer. The number of times an examinee can make a particular error type ranges from five on Type D to 38 on Type B.

Each subject in the sample received 10 pretest scores and 10 posttest scores, with each score representing the total number of items on which the subject made an error of a given type. The 10 scores were then submitted to a principal components analysis with 1.00 in the diagonal, followed by a varimax rotation with a three factor solution. This procedure was carried out initially on a combined sample of 1147 tests, including pretests of normal children, pretests of trained EMR children, posttests of trained normal subjects, posttests of non-trained normal subjects, and posttests of trained EMR subjects. As will be shown in the Results section, neither the principal components solution nor the varimax rotation yielded factor structures that could be meaningfully interpreted. For this reason, the same analysis was performed on these five sets of tests separately. (Because the sample of untrained EMR subjects was only 46, the tests of this group were not factor analyzed.)

Results

Table 2 presents the mean frequencies in percentages with which each group made the 10 types of errors. These percentages were calculated by dividing the frequency for each error by the number of errors of that type, contained in all 36 test items, and multiplying the quotient by 100. Since the number of times it is possible to make each type of error varied, this procedure was used to permit comparison of frequencies on one error type with another. For example, on the pretest (R1), trained EMR subjects made 13.49% of all possible Type F errors and 26.62% of the possible Type G errors.

The table confirms the finding obtained with another sample (Budoff & Corman, 1975) which showed that both EMR and normal children had a higher percentage of Type G and H errors than any other error type on both pretests and posttests. These errors involve pattern repetition immediately above the space to be filled (Type G) and immediately to the left of the spcce to be filled (Type H). It is clear that frequencies for these two errors decreased markedly with training for both groups of subjects. Percentages of Error Types A, B, and J tended to be the smallest for all groups of subjects regardless of test administration.

In the principal components analysis, the first three principal components together accounted for 53.2% of the total variance for the combined sample of 1147 tests. Latent roots were 3.051, 1.294, and .980, repsectively. Since the third latent root fell just short of 1.00, the three factor solution was considered appropriate according to Kaiser's criteria (roots greater than 1.00). Of the total variance, 30.5% was contributed by Factor I, 12.9% by Factor II, and 9.8% by Factor III.

According to Raven's scheme (Table 1), four factors might have been anticipated: errors involving a difference (Types A and B), inadequate individuation (Types C, D, and E), pattern repetition (Types F, G, and H), and incomplete correlates (Types I and J). On the whole, the error pattern revealed by the three principal components with the combined test sample did not seem logically interpretable according to Raven's description of error types. Error types which had high loadings (>.50) on the first principal component were B, C, D, E, F, I, and J. On the second principal component Error Types F, G, and H had high positive loadings (>.40), and A had a negative loading of -.446. Error Types A and H had negative loadings above .40 on the third principal component as well.

Table 2

Mean Percent Frequency of Ten Types of Errors

Error type	Total possible (percentage base)	EMR				Normal			
		Trained		Not trained		Trained		Not trained	
		R1	R2	R1	R2	R1	R2	R1	R2
A	11	.52	.37	.61	.40	.12	.13	.19	.00
B	38	2.28	1.63	2.11	2.93	.62	.20	.31	.28
C	22	5.64	4.19	7.07	5.86	2.65	2.04	2.71	1.79
D	5	4.92	4.90	5.78	4.88	3.62	1.24	2.76	3.10
E	12	7.26	5.32	6.85	6.67	4.31	1.15	2.58	1.19
F	13	13.49	11.44	13.68	10.08	8.36	6.37	8.18	6.41
G	20	26.62	21.57	27.22	26.11	21.31	10.86	17.40	12.09
H	19	20.53	16.82	22.57	24.79	19.39	10.45	17.42	14.54
I	14	8.36	6.92	7.78	9.36	5.32	3.69	5.24	4.25
J	26	3.66	2.75	4.53	3.93	2.30	1.27	1.60	1.56
N		350	318	45	45	149	145	94	84

Table 3 presents the types of errors with high loadings on the three rotated factors for the combined sample of tests and for the five separate sets of tests. In the five separate analyses, the three factors obtained accounted for 34% to 42% of the total variance. It is clear that for the combined sample of 1147 tests, the error types with high loadings on the rotated factors were not the same as those with high loadings on the three principal components. Varimax rotations with the five sets of tests did not appear to reveal a marked similarity of error patterns among them, although Types A, B, and J had high loadings on Factor I in most of the analyses. Moreover, the separate factor structures did not seem to lend themselves to any more meaningful interpretation than was possible for the combined sample of tests.

Discussion

Raven has pointed out that qualitative comparisons of error responses on this test have been relatively unsuccessful in the past. He stated, "This has been due partly to the unequal distribution of erroneous figures in the scale as a whole, and partly to unsuccessful attempts to classify erroneous choices in accordance with their character qualities." (Raven, 1965, p. 10.)

For this sample factor analysis did not reveal patterns of error types on the Raven Progressive Matrices (Sets A, AB, B) which could be meaningfully interpreted. In addition, factor analysis did not serve as powerful means of data reduction for the five separate sets of tests: the three factors accounted for only 34% to 42% of the total variance in the five analyses. In a 10 X 10 correlation matrix one might expect three variables to account for 30% of the total variance, since each

Table 3

Error Types with Factor Loadings Greater than .30 on Varimax Rotations

	Combined test sample	Pretests of normals	Posttests of trained normals	Posttests of untrained normals*	Pretests of trained EMRs	Posttests of untrained EMRs
Number of tests	1147	243	147	84	353	319
Factor I	A B J	B C D E F J	A B E J	B C E	A B D J	A B J
Percent variance	30.5	19.9	22.0	22.3	16.7	14.0
Factor II	C E F G H I J	G	C F G I	H I	B C E F I	C D E
Percent variance	12.9	8.0	11.1	10.0	13.5	10.5
Factor III	D H	E F H I	G H	C F G J	G H	E F G H I
Percent variance	9.8	6.4	9.1	8.3	6.5	9.6
Percent total variance	53.2	34.3	42.2	40.6	36.7	34.1

*No subject in this sample made any Type A errors.

variable would contribute one-tenth of the variance. Factor III accounted for less than one tenth of the variance in all analyses. Only Factor I consistently contributed more than 10% in every analysis. This factor is composed largely of errors which involve no figures (A), irrelevant figures (B), and incomplete correlates (J) (i.e., nearly correct), a pattern which does not lend itself to logical interpretation. It may be that they load on this factor consistently because these error types occur most consistently with the lowest incidences for the average and low IQ students prior to or following problem-relevant training.

References

Budoff, M., & Corman, L. Effectiveness of learning potential training on reduction of errors on Raven's Progressive Matrices. Studies in Learning Potential, 1975, No. 48.

Budoff, M., Corman, L., & Litzinger, S. A manual for the Raven Learning Potential test. Studies in Learning Potential, 1975, No. 59.

Raven, J. C. Guide to using the Coloured Progressive Matrices. London: H. K. Lewis, 1965.

Footnote

[1] This research was supported by Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, Office of Education, NE-G-00-3-0016 from the National Institute of Education, and MH R01-18553 from the National Institute of Mental Health, all under the Department of Health, Education and Welfare.

Abstract

This study sought to determine the effectiveness of learning potential training on improving children's problem-solving skills on Raven's Progressive Matrices. A sample of 553 retarded and nonretarded students were pretested, randomly assigned to trained or nontrained groups, and posttested. Analyses of variance on factor scores derived from a varimax rotation of item responses revealed that learning potential training was effective in increasing the ability to reason by analogy of both retarded and nonretarded children. The learning potential procedure provided learning experiences related to reasoning skills which may not have been previously acquired in school.

Effectiveness of a Learning Potential Procedure in Improving Problem-Solving Skills of Retarded and Nonretarded Children

The learning potential (LP) assessment procedure was developed by Budoff and his associates as an alternative to traditionally verbally-biased IQ tests for estimating general ability to reason. This assessment approach is based on a conceptualization of intelligence which stresses trainability, or the ability to profit from learning experiences. It assumes that the prior and continuing experiences of poor and/or nonwhite children do not allow them to spontaneously acquire school-relevant skills in their natural environment. The learning potential paradigm provides subjects with training experiences directly relevant to the reasoning task presented, so they can apply, in microcosm, their problem-solving ability and show whether they can improve their performance on the task. Ability to profit from LP training is not as dependent on demographic factors (such as race and social class) which are usually correlated with IQ (Budoff & Corman, 1974). LP assessment replaces the traditional product-oriented test with a three-stage procedure which includes a pretest, two training sessions, and a posttest.

Previous research conducted with the intellectually average sample employed in this study ($\underline{N} = 379$) has revealed that LP training was not differentially effective by race (Budoff & Corman, Note 1). Middle class children in this sample tended to increase their Raven scores to a greater extent than lower class children following training; however, 40% of the nonretarded lower and lower middle class students attained posttraining scores as high or higher than the pretest mean of the middle class students. Before training, only 15% of the lower class students exceeded the middle class pretest mean. Many children from non-middle class homes, who might not have had frequent access to appropriate school-preparatory experiences and who tend to score poorly on traditional IQ tests, were able to demonstrate their nonverbal reasoning ability after training on similar tasks was provided.

Previous research with the normal and educable mentally retarded (EMR) sample of this study (N = 553) sought to determine the effectiveness of the LP training procedure in improving the problem-solving ability of these subjects, as reflected by their total scores on Raven's (1956) Coloured Progressive Matrices (Budoff & Corman, Note 2). The study indicated that trained subjects improved their overall performance on the nonverbal reasoning tasks to a greater extent than nontrained subjects, and that training relevant to the problem types presented in the Raven Progressive Matrices was equally effective for nonretarded and retarded subjects. Similarly, the Raven's LP procedure was also shown to reduce the frequency of problem-relevant errors of normal and retarded students on the post-training administration (Budoff & Corman, Note 2).

The aim of this study was to determine the particular subskills measured by the Raven test for which learning potential training is most effective with young intellectually normal and educable mentally retarded children. While it is known that training procedures improve total scores on the Raven test, it is not knownwhether all Raven subskills are equally affected by training, or whether training affects some subskills and not others. Previous work on the Raven (Corman & Budoff, 1974) has established an invariant four-factor structure for Sets A, AB, B of the Raven's: I Continuity and Reconstruction of Simple and Complex Structures, II Discrete Pattern Completions, III Reasoning by Analogy, IV Simple Continuous Pattern Completions. The purpose of the present study was to assess impact of training on each of these skills.

Method

Subjects

The sample consisted of 379 normal and 174 special class students who attended schools in a large urban community in upstate New York. Normal subjects ranged in age from six to eleven and were evenly distributed among the first through fifth grades. Fifty-three percent of these subjects were black and 28% of the normal students' fathers were employed in business or professional occupations. The majority were blue collar or clerical workers. EMR subjects ranged in age from 5½ to 14 years, with a mean age of 10 (SD = 2). All but 5% were from low socioeconomic backgrounds (welfare, blue collar, or clerical) as indicated by father's occupation, and 43% were black. The mean IQ of the EMR subjects on the WISC or Stanford-Binet was 68.4, with a standard deviation of 9.0. Fifty-five percent of both normal and EMR subjects were male.

Procedure

Both normal and EMR subjects were assigned to trained and nontrained groups with a randomized block procedure at the outset of the study. On the basis of pretest scores on the Coloured Progressive Matrices (Raven, 1956), normal subjects were randomly assigned to a nontrained group or to a group which received LP training in groups of six to thirteen students. EMR subjects were randomly assigned to one of three groups on the basis of their pretest scores: a nontrained group, a group which received individual training, and a group which received training in groups of five to eight students. Since analysis of variance revealed no significant difference between posttest scores of EMR subjects who were trained individually or in a group, subjects who received both types of training were combined to constitute the trained group of EMR children.

A total of 1089 tests were administered to children in classroom-size groups. Pretests were given to the total sample of 379 normal and 174 EMR children, and posttests (Coloured Progressive Matrices) were administered to 172 trained and 199 nontrained normal subjects and to 110 trained and 55 nontrained EMR subjects. All posttests were administered within two weeks after pretests had been given.

Each child in the training groups was given a training booklet which contained nontest problems dealing with pattern completion, orientation of elements within a pattern and double classification problems. The trainer presented the problems on 2 X 2 slides from a Kodak Carousel Projector on a blackboard. The students were required to draw in the missing element for the design and mark the appropriate choice from the six choices presented on the lower half of the page. For the double classification problems, it was found that children could easily derive one attribute at a time but often did not hold the first attribute in mind while they derived the second relevant attribute. During development of the training procedure, the child's understanding was facilitated by having him draw the relevant attributes, one at a time, as he derived them. This helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems mentally with little trouble.

The requirements of each problem type were presented in meaningful designs initially, e.g., an American flag with a piece missing, and then a geometric form to attune the child to the basic format of the Matrices test problems. Individual children were called to indicate the correct choice, and to give reasons for their choice. A slide with the answer included allowed them to compare their choice and to correct it, if necessary. (See Budoff & Litzinger, Note 3, for details of test administration and training procedure.)

Results

After responses to the 36 items on the 1089 tests had been submitted to a principal components analysis with 1.00 in the diagonal, followed by a varimax rotation, four factor scores based on this rotation were computed for each subject. Corman and Budoff (Note 4) have demonstrated the comparability of factor structures of normal and EMR subjects on pretests and posttests. Repeated measures analyses of variance on 533 subjects tested and the effects of IQ Group (EMR versus normal), Training Group, and Test Session on factor scores obtained with the four rotated factors. In these four analyses, the effect of the interaction between Training Group and Test Session was the most critical effect, testing differences from pretest to posttest due to training. The Three-way interaction of these two factors and Group permitted examination of differential effects of training on normal and EMR subjects.

Results of these analyses are presented in Table 1 and cell means are shown in Table 2. The factor score means can be interpreted as standard score units, since the factor scores for each subject were calculated to have a mean of zero and standard deviation of one.

Insert Tables 1 and 2 about here

A significant main effect for IQ Group was found on Factors I (Continuity and Reconstruction of Simple and Complex Structures) and III (Reasoning by Analogy). Mean scores (Table 2) indicated that normal children performed better than EMR children on these factors. The fact that normal and EMR subjects did not significantly differ on Factors II (Discrete Pattern Completions) and IV (Simple Continuous Pattern Completion) reflected the presence of a ceiling effect on these factors. Item analysis revealed that items with high loadings on Factors II and IV were the easiest items on the pretest. Since most students answered these items correctly on the pretest, the ceiling effect restricted the range for demonstrating improvement on these two factors. from the pretest to the posttest.

The main effect for Training Group was significant only on Factor III (Reasoning by Analogy), on which the trained subjects scored higher. The IQ Group X Training Group interaction on Factor II was significant because of the greater discrepancy between scores of trained and nontrained EMR children, relative to normal children, on this factor. That is, when pre- and posttest scores are averaged, trained EMR children perform better than nontrained EMR children and this difference is greater than that between trained and nontrained normal subjects (Table 2). The three-way interaction (Group X Training Group X Test Session), which would indicate differential effects of training on the two IQ groups, was not significant on any of the four factors.

The critical Training Group X Test Session interaction was found to be significant on Factor III: trained subjects improved more from pretest to posttest than nontrained subjects on Reasoning by Analogy problems. Since the three-way interaction was not significant, the training effect on this skill was consistent for both EMR and normal children. The significant IQ Group X Session interaction on Factor III indicated that when training is not taken into account, normal children improved more than EMR subjects on this factor.

Means in Table 2 indicate that on Factor I (Continuity and Reconstruction of Simple and Complex Structures) trained normal and EMR subjects gained more than nontrained subjects, but the difference from pretest to posttest was less marked than on Factor III and was found to be nonsignificant. Relative to the nontrained groups, the greatest increment of the trained groups (both normal and EMR) was made on Factor III, Reasoning by Analogy.

Discussion

Learning potential training has been found to be an effective means of improving the problem-solving ability of both retarded and nonretarded children. Embedding training within the assessment procedure and relying on the posttraining score as the primary predictor, indicates fewer differences between IQ defined EMRs and nonEMRs from low income homes in profitting from suitably structured experiences relating to the reasoning problems. Results of this study indicated that, of the four subskills measured by the Raven Progressive Matrices, training was most effective in improving the ability to reason by analogy of both EMR and young normal children, i.e., in changing these students' understanding of double classification problems. Changes in two of the factors which involve simple perceptual patterns were not obtained because most of the children solved them on the pretest. The fourth factor was composed of a motley collection of medium difficulty to complex perceptual and orientation problems that demonstrated changes following training that were in the predicted direction for retarded subjects, though the differences were not significant.

Results of this study lend further support to previous investigations of the learning potential procedure, which have consistently indicated that students in special and regular classes are able to profit from learning experiences that exposed them to problem-solving strategies related to reasoning tasks.

Since the primary effect of the training is on the Reasoning by Analogy Factor items, providing appropriate experiences does improve these low income children's Level II abilities in a fashion parallel to that demonstrated for Level I simple learning tasks. The primary condition required is that the corrective feedback opportunities present in repeated trials learning experiments also be made available to the students from low income backgrounds. Training embedded within the assessment procedure provides this condition.

Reference Notes

1. Budoff, M., & Corman, L. The effectiveness of a group training procedure on the Raven learning potential measure with children from diverse racial and socioeconomic backgrounds. RIEP-print No. 58, 1973.
2. Budoff, M., & Corman, L. Effectiveness of learning potential training on reduction of errors on Raven's Progressive Matrices. RIEP-print No. 48, 1975.
3. Budoff, M., & Litzinger, C. S. A manual for the Raven learning potential test. RIEP-print No. 59, 1975.
4. Corman, L., & Budoff, M. Factorial composition of the Raven Progressive Matrices in the context of learning potential measurement. RIEP-print No. 62, 1975.

References

Budoff, M., & Corman, L. Demographic and psychometric factors related to improved performance on the Kohs Learning Potential procedure. American Journal of Mental Deficiency, 1974, 78, 578-585.

Corman, L. & Budoff, M. Factor structures of retarded and nonretarded children on Raven Progressive Matrices. Educational and Psychological Measurement, 1974, 34, 407-412.

Raven, J. C. Coloured progressive matrices: Sets A, AB, B. London: H. K. Lewis, 1956.

Footnote

This research was supported by grants OEG-0-8-080506-4597 from the U. S. Office of Education, NE-G-00-3-0016 from the National Institute of Education, and MH R01-18553 from the National Institute of Mental Health, all under the Department of Health, Education and Welfare. The Research Institute for Educational Problems is located at 29 Ware Street, Cambridge, Ma., 02138.

Footnote

This research was supported by grants OEG-0-8-080506-4597 from the U. S. Office of Education, NE-G-00-3-0016 from the National Institute of Education, and MH R01-18553 from the National Institute of Mental Health, all under the Department of Health, Education and Welfare. The Research Institute for Educational Problems is located at 29 Ware Street, Cambridge, Ma., 02138.

Table 1

F Ratios from Repeated Measures Analyses of Variance on Factor Scores

Source	Factor I[a]	Factor II[a]	Factor III[a]	Factor IV[a]	df[b]
IQ Group (EMR vs. normal)	40.455***	1.915	20.019***	.128	1,529
Training Group	.350	2.610	3.917*	1.569	1,529
IQ Group X Training Group	.049	5.544*	1.009	.006	1,529
Session	44.069***	.004	27.652***	1.969	1,529
IQ Group X Session	1.438	1.116	4.823*	1.223	1,529
Training Group X Session	1.720	2.089	13.964***	.063	1,529
IQ Group X Training Group X Session	.294	negligible	.680	1.073	1,529

[a] Factor I--Continuity and Reconstruction of Simple and Complex Structures; Factor II--Discrete Pattern Completions; Factor III--Reasoning by Analogy; Factor IV--Simple Continuous Pattern Completions.

[b] 20 subjects were deleted because of missing posttest data.

*p <.05

**p <.01

***p <.001

Table 2

Mean Factor Scores of Trained and Nontrained EMR and Normal Subjects on the Pretest and Posttest

	Factor I[a]		Factor II[a]		Factor III[a]		Factor IV[a]		
	pre	post	pre	post	pre	post	pre	post	N
Normal									
Trained	.085	.301	-.076	.061	-.040	.518	.001	-.065	173
Nontrained	.079	.240	.070	.038	-.062	.062	-.003	.105	206
EMR									
Trained	-.539	-.161	.008	.078	-.325	-.075	-.149	.041	114
Nontrained	-.559	-.300	-.187	-.316	-.298	-.289	.008	.089	60

[a] Factor I--Continuity and Reconstruction of Simple and Complex Structures; Factor II--Discrete Pattern Completions; Factor III--Reasoning by Analogy; Factor IV--Simple Continuous Pattern Completions.

PLEASE REFER TO ED 108 439

STUDIES IN LEARNING POTENTIAL

THE PICTURE WORD GAME: A NONVERBAL TEST OF THE ABILITY TO USE LANGUAGE-RELATED SYMBOLS

Louise Corman and Milton Budoff

Volume 4, Number 77
1974

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

PLEASE REFER TO ED 108 439

THE PICTURE WORD GAME: A NONVERBAL TEST OF THE ABILITY TO USE LANGUAGE-RELATED SYMBOLS[1]

Louise Corman and Milton Budoff
Research Institute for Educational Problems

Budoff and his associates have developed a training-based procedure for assessing reasoning ability on nonverbal tasks (Budoff, 1970). In this assessment strategy training assumes a critical role, particularly for the child from a poor and/or nonwhite background who may learn different cognitive strategies in expressive formats other than those presumed to be available by traditional tests. The training helps the child to narrow the cognitive gap between his previously learned problem-solving strategies and those implicit to the problems he must ordinarily solve on the middle-class-biased tests he encounters. Inclusion of training in the assessment procedure also minimizes the artificiality of the test situation. Repeated contacts with the materials in a context of support and teaching allow the school-failing child to develop a sense that he can be competent. Without this competence boost, he tends not to perform at his best, implicitly expecting failure (Zigler, 1966)

The essence of this assessment strategy, then, is to impose some control on the potentially negative effects on the child's test performance of prior life experiences. Two types of effects can be considered: those which are due to problem solving experiences which differ from the abstracting verbal conceptual types of skills expected of school children, and the negative effects of failing to perform well on the tests the child has taken during his school years. Test scores after training should reflect the child's ability under optimized conditions in which he is familiar with the task and its demands, has had success in solving problems similar to those on the test, and has had the opportunity to learn and apply relevant strategies.

The goal of this study was to apply the rationale from training based assessment to tasks that are language related, since previous efforts had focused on nonverbal reasoning tasks. The specific aim was to develop a procedure that would permit the low SES, low school achieving child to show his/her proficiency in utilizing language related symbols, following a systematic learning experience. In this task, the child learns a new symbol system for common words that he understands, and must apply this understanding in generating or decoding sentences. A critical feature of this language measure is that it not make excessive demands on the previously developed verbal skills of students who may have experienced difficulty with verbal tasks.

The Picture Word Game (PWG) was conceived as a modification of the Semantic Test of Intelligence (STI) constructed by Philip Rulon. The STI is a language-related measure in which the examinee learns to associate a geometric symbol with a pictured object or action (man, woman, runs, pushes). The student must "read" the sequence of symbols and choose the appropriate picture from a multiple choice array. The task is an analogue to reading the words in a sentence and indicating the picture that best represents the meaning of the sentence. As such, it was thought that the task might represent a means by which one could examine verbally related competencies of children in a training-based format, without the necessity of verbal expressive materials or reading words, skills in which these children are often deficient.

The STI is administered as a timed test and consists of 217 items, including 109 items with one symbol, 49 with two symbols, 36 with three symbols, and 23 with four symbols. The symbols for each noun and verb are introduced as single symbols, and defined by accompanying pictures in a multiple choice format on tuition pages which are not scored. Tuition pages are used to introduce 2-, 3-, and 4-symbol "sentences." Instructions are pantomimed. The symbols and the pictures which define their meaning are presented on each double page so that memory for the meaning of the symbol is not a factor influencing performance.

The STI was developed as a measure of military trainability for illiterate recruits who had failed the literacy requirement for entry into the Marine Corps. Validity of the test as a language measure was evidenced by Rulon and Schweiker's (1953) finding that recruits who did well on the STI also successfully completed a literacy course to meet the eligibility requirements of the Marine Corps.

In a previous study (Gimon, Budoff, & Corman, 1974), the investigators administered the STI to 76 Spanish-speaking children from low income families, who ranged in age from 6 to 13 years. Results indicated that a ceiling effect occurred with children over 8 years old. The mean number of one-symbol items passed by over 50% of the total sample was 92 of 109 (84%), and all but seven of the 49 two-symbol items were passed by more than half the children. These findings revealed that the majority of items on the test were too easy for the children in this age range. This investigation also revealed that STI scores were significantly correlated with vocabulary scores on both the Spanish and English versions of the WISC. The investigators concluded from these results that the tasks used in the STI were related to verbal ability, but that the difficulty level of the test would have to be broadened in order for it to be used with normal children spanning a wider age range.

The Picture Word Game was constructed to provide a minimally verbal measure of language ability of children in the first through fifth grades. Development of the test was conducted in three phases: i) construction of the original test and accompanying training procedure, ii) test development, i.e., revision of the original test, and iii) evaluation of psychometric characteristics of the final test.

Construction of the Original Test

The following principles represent departures from the STI format and were used with the Picture Word Game in an effort to broaden the difficulty level of the STI:

1. The Picture Word Game is a power test, administered as an untimed measure. Eliminating the effect of speed in performance was considered desirable to reduce spurious inflation of variability and consequently reliability.

2. The vocabulary in the test, i.e., the number of symbols, was increased from nine "words" in the STI to 16 in the Picture Word Game.

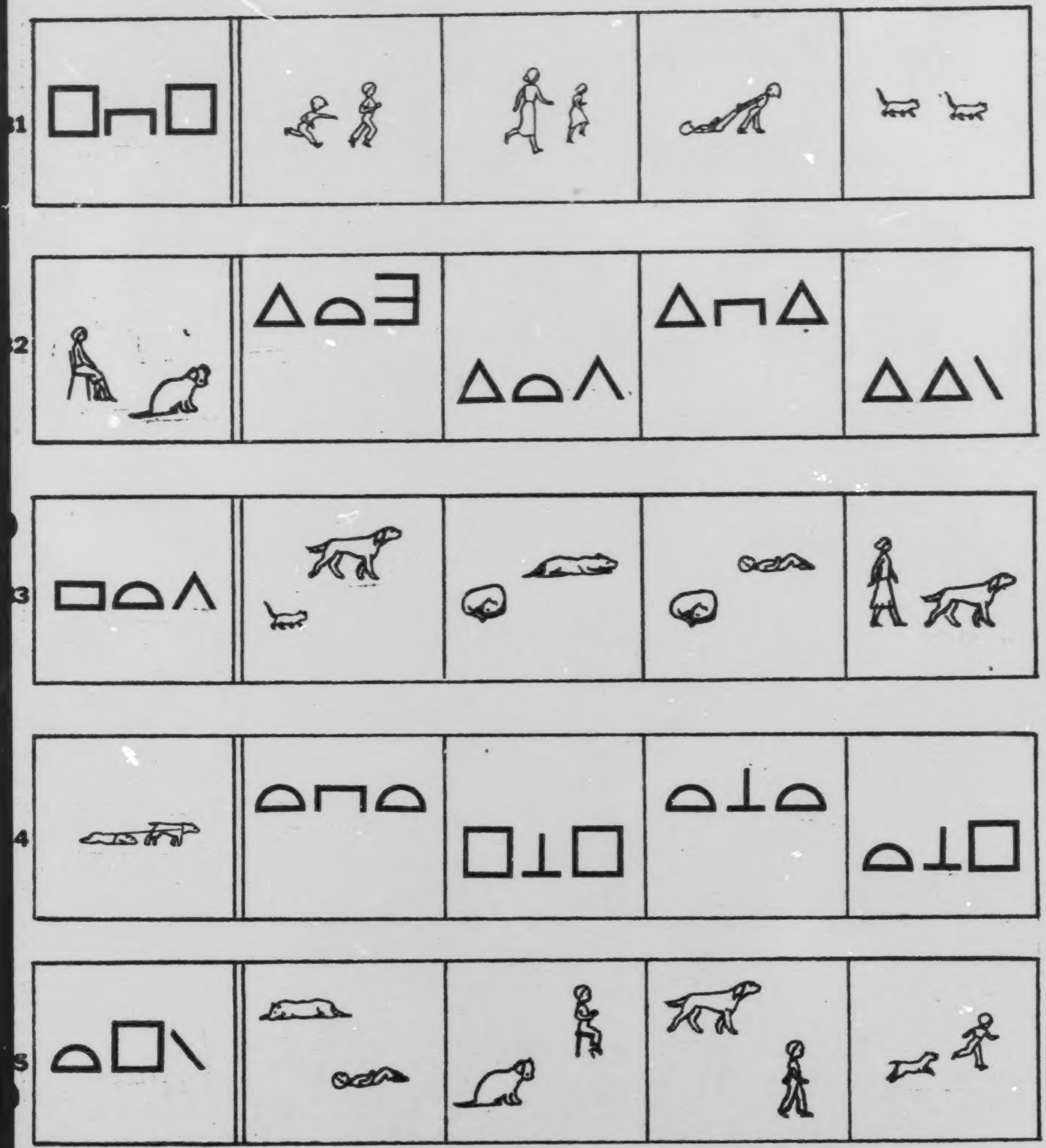
3. The concept of symbols representing numbers was included, in addition to noun and verb representation.

4. Whereas all items in the STI require translation from symbol to picture, approximately half the items in the Picture Word Game require translation from picture to symbol. The investigators believed that translation from picture to symbol would be a more difficult task, especially when three or more symbols were to be read. An internal measure of the effectiveness of the teaching sequence would be the extent to which the child can utilize the symbol in both directions--to decode the symbols into their pictorial equivalents, and to generate sentences in the "new" language that would explain the picture. (See Figure 1).

5. While the STI contained 217 items, including a large number of very easy one and two symbol items, the original version of the Picture Word Game contained 60 items with a much smaller proportion of one and two symbol items. The intent was to reduce the number of items on the final form even further so that the Picture Word Game could be administered in one class period.

6. While the most difficult items on the STI require reading a "sentence" of four symbols, the Picture Word Game included items of five symbols as well. It was not possible to construct a meaningful sentence of more than five symbols within this format, when these symbols represented only nouns, verbs, and numbers.

FIGURE 1



The original form of the Picture Word Game contained 60 items, with odd-numbered items requiring translation from symbol to picture and even-numbered items requiring translation from picture to symbol. The test included six one-symbol, 12 two-symbol, 18 three-symbol, 18 four-symbol, and 6 five-symbol items. Vocabulary comprised 16 symbols: five nouns (cat, dog, woman, boy, horse), five intransitive verbs (sit, lie down, walk, stand, run), three transitive verbs (pull, chase, carry), and three numbers (two, three, and four).

A training procedure was developed such that training could be entirely concluded before administration of the test, rather than embedding training within the test as in the STI. Oral instructions were used in the training, instead of relying solely on pantomimed instructions. The latter procedure used in the STI was considered artificial by the investigators on the basis of a previous administration of the STI (Gimon, Budoff, & Corman, 1974).

A training booklet was devised which contained 10 of the 16 symbols used in the vocabulary of the test. A page with all the symbols in the vocabulary was attached to the back of the training booklet, so that it could be easily detached and referred to by the student during the training session. The six symbols contained in the test but excluded from training were one noun, one intransitive verb, one transitive verb, and the three numbers. The training booklet contained items similar to those in the test, and slides were made of each training item. The trainer simultaneously displayed the slides while explaninig the principles of the tasks and techniques for solving them to students using the training booklet.

Test Revision

Subjects and Procedure

Pilot testing was conducted in June of 1973, with a sample of 205 students from a low income school district in an urban community in Massachusetts. The subjects constituted ten first through fifth grade classrooms, two classes per grade, and were evenly divided by sex. The mean age of the sample was 9 years, 2 months, with a standard deviation of 1 year, 7 months.

All students were trained in one class period and tested in class size groups in a second class period three days following training. A student's score was calculated as the percent of items he answered correctly of the total number of test items.

Results

The test showed a high degree of internal consistency reflected in a KR20 reliability coefficient of .95. The mean percent of correct responses for the total sample was 75.5 (SD = 20.8). The mean percent correct for students in grades one to five, respectively, were: 62.7 (±21.3), 63.0 (±21.2), 86.9 (±11.8), 79.5 (±11.8), and 89.1 (±11.8). These figures indicated that, despite efforts to make this test more difficult than the STI, a ceiling effect was found for students at or beyond the third grade. Mean scores by grade level reflected a dichotomy between first and second graders on the one hand and third through fifth graders.

Selection of Items for the Final Test

Difficulty levels and discrimination indices were obtained for each item with the total sample and subjects in each grade. Items were selected for the final test which met three criteria: i) the difficulty level of test items for the total sample was evenly distributed throughout the test, i.e., approximately equal numbers of easy and moderate-to-hard items were retained, ii) to the extent it was possible, each item selected showed a gradual increase in difficulty from grades one to five, and iii) the discrimination of each item was not less than .25 for the total sample. These procedures were used in an attempt to produce test scores that would reflect a developmental trend and reduce the likelihood of a ceiling effect on the final test. The revised test consisted of 37 items, including one one-symbol, 7 two-symbol, 9 three-symbol, 9 four-symbol, and 11 five-symbol items.

Other revisions based on pilot testing consisted of improvements in pictures used in specific items. An effort was made in both the test and training booklets to sharpen pictures which the children had difficulty interpreting. Appendix A contains the final training instructions and student's training booklet. The final version of the test is presented in Appendix B.

Characteristics of the Final Test

Subjects

The sample for this phase of the study, which was conducted in January, 1974, consisted of 90 students from a low income area in Massachusetts. Subjects were evenly divided into five first through fifth grade classrooms, one class per grade, and attended an elementary school in the same urban community as the school from which the pilot sample was drawn. Subjects were evenly divided by sex. The mean age of the sample was 8 years, 10 months (SD = 1 year, 9 months).

Procedure

All students were trained in one class period and tested in class size groups within two days following training. In addition, all students' scores on the Stanford Paragraph

Meaning and Vocabulary Subtests were obtained. The Primary II level of this test had been administered to third through fifth graders by the school when these students were in the third grade, and scores were obtained from their school records. First and second graders were group administered the Primary I level after Picture Word testing had been completed.

Results

Discrimination and difficulty levels of final test items are presented in Table 1. Discrimination indices of almost all items were high, and this fact was reflected in the high KR20 reliability coefficient of .93. Mean item difficulties revealed that several procedures used to modify the STI were successful in increasing the difficulty level of the Picture Word Game when the total sample is considered: a) items requiring translation from symbol to picture were significantly harder than items which required translation from picture to symbol ($F = 63.64$, 1/85 df, $p < .001$); b) the difficulty of items increased linearly as the number of symbols increased from two to five ($F = 70.27$, 3/255 df, $p < .01$), and c) the two most difficult sets of items on the test were items employing the concept of numbers which had not been taught, and all of which had three or more symbols (mean = 59.3%) and the five symbol items (mean = 53.4%). The STI does not include either of these item types.

Despite the effectiveness of these procedures, however, mean total scores (in terms of percent correct) for each grade again revealed a ceiling effect for children beyond the third grade whose average score exceeded 30% correct (Table 2). Total score means reflected the same dichotomy between grades 1 to 2 and 3 through 5 which was found on the pilot test, despite item selection procedures which retained items with the most marked linear developmental trend. The table indicates that the gap between second and third graders widened as the number of symbols increased from three to four. Items involving number concepts and those requiring translation from picture to symbol also differentiated first and second graders from the rest of the sample.

Mean stanines on the Stanford Comprehension and Vocabulary subtests were relatively comparable among the five grades and revealed that these children's reading skills were in the moderate to low range in relation to national norms. The mean stanine for the total sample was 4.3 (±1.8) on comprehension and 4.1 (±1.4) on vocabulary. Scores on the Picture Word game were correlated with scores on these two subtests to provide evidence of validity of the Picture Word Game as a language measure. Coefficients of .37 and .34 were obtained with comprehension and vocabulary, respectively. While these coefficients

TABLE 1

Discrimination and Difficulty Levels of Final Picture Word Game Items

Item	Number of Symbols	Stimulus Mode[a]	Number Concept?	Discrimination	Difficulty
1	1	P		.04	.96
2	2	S		.29	.90
3	2	P		.45	.81
4	2	S		.32	.92
5	2	S		.27	.92
6	2	P		.43	.84
7	2	S		.50	.92
8	2	S		.44	.90
9	3	P		.44	.79
10	3	S		.46	.79
11	3	P		.49	.62
12	3	S		.57	.88
13	3	P		.73	.80
14	3	S	yes	.56	.61
15	4	P	yes	.44	.50
16	3	S	yes	.65	.58
17	3	P	yes	.47	.63
18	3	S		.62	.87
19	4	P		.47	.77
20	4	S	yes	.51	.65
21	4	P	yes	.46	.44
22	4	S	yes	.72	.77
23	4	P		.39	.63
24	4	S	yes	.74	.77
25	4	P	yes	.59	.68
26	4	P		.62	.63
27	5	P	yes	.67	.58
28	5	S		.59	.67
29	5	P		.65	.70
30	5	S	yes	.74	.61
31	5	P	yes	.62	.41
32	5	S	yes	.67	.64
33	5	P		.62	.38
34	5	S	yes	.75	.58
35	5	P		.58	.43
36	5	S	yes	.57	.43
37	5	P		.58	.44

[a] Items marked "P" require translation from picture to symbol; items marked "S" require translation from symbol to picture.

TABLE 2

Mean Percent Correct by Grade on the Picture Word Game

	Total Test		Number of Symbols								Stimulus Mode				Number Concept		N
			Two[a]		Three		Four		Five		Picture to Symbol		Symbol to Picture				
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	X	SD	X	SD	X	SD	N
Total Sample	68.8	23.4	89.7	16.2	73.0	26.7	64.9	29.2	53.4	35.0	63.5	24.5	74.5	24.1	59.3	30.5	90
Grade 1	45.0	21.1	74.4	23.2	44.4	30.2	46.1	32.4	23.2	24.5	40.3	20.5	50.0	24.7	33.0	23.6	20
Grade 2	55.2	19.5	92.3	11.9	65.1	23.1	48.7	33.2	25.5	26.9	47.4	21.1	63.5	19.8	37.1	25.9	21
Grade 3	80.4	11.5	93.3	9.0	85.2	11.2	71.9	9.8	73.9	23.0	73.7	13.0	87.4	14.0	75.6	20.0	15
Grade 4	84.5	10.0	95.3	10.7	86.1	8.3	84.0	13.6	75.6	21.7	81.9	9.8	87.2	11.7	77.9	16.6	16
Grade 5	87.7	8.5	95.8	5.9	92.0	12.2	82.1	14.4	82.8	13.2	83.0	11.2	92.6	7.6	84.1	13.9	18
Number of Items	37		8		9		9		11		19		18		15		

[a] Item 1 with one symbol was included in this group.

are low in terms of concurrent validity, they are similar in magnitude to the validity coefficient of .385 obtained by Rulon and Schweiker (1953) with the STI.

Discussion

A training-based procedure clearly has advantages for assessing language skills of low SES children. These children typically have a verbal deficit which inhibits them from displaying their maximal competence on traditional verbal tests. Training provides a supportive context which can permit them to demonstrate optimum performance.

The basic principles employed in the STI would appear to lend themselves to adaptation for use as a language measure with children of a broad age range. Although certain test construction procedures were effective in increasing the difficulty of the STI, attempts to broaden the difficulty level of the test as a whole were largely unsuccessful. Like the STI, the Picture Word Game was shown to be too easy for children beyond the third grade, despite the fact that comprehension and vocabularly skills of this sample were below average on national norms.

In its present form the Picture Word Game appears to be most useful with second graders. It could be argued that the training session was so beneficial in maximizing test performance that training was partially responsible for the ceiling that resulted in the test. Some might suggest that training only items which require translation from symbol to picture might lower the test ceiling and at the same time permit examination of transfer of training to the more difficult picture to symbol items on the test. It is doubtful, however, that this procedure would effectively reduce the test ceiling; despite the fact that items involving number concepts or five symbols were excluded from the training, children beyond the third grade had little difficulty in translating such items on the test.

Results of this study indicated that the Picture Word Game measures a unitary ability which is related to language skills. It is likely that facility in translation required by this test is mastered by children of normal intelligence after the third grade and that, within the STI format, quantitative modifications are not sufficient to tap the increasingly complex language skills learned by children in the intermediate grades. Use of the Picture Word Game in further research with educable mental retarded children might prove to be quite fruitful; it is unlikely that the ceiling effect would be evident with these children for whom a minimally verbal, training-based measure is clearly appropriate and could serve a critical need.

References

Budoff, M. Learning potential hypothesis: Its effects on teacher expectancies for disadvantaged students. Paper presented at the Conference on Teacher Expectancies, Georgia State University, Atlanta, Georgia, October, 1970.

Gimon, A., Budoff, M., & Corman, L. Applicability of learning potential measurement with Spanish-speaking youth as an alternative to IQ. Final Report, 1974. Grant OEG-1-72-0020 (509), Office of Education.

Rulon, P., & Schweiker, R. Validation of a nonverbal test of military trainability. Final Report, June, 1953, Department of the Army.

Zigler, E. Mental retardation: Current issues and approaches. In L. W. Hoffman & M. L. Hoffman (Eds.), Review of child development research, Vol. 2. New York: Russell Sage Foundation, 1966.

Footnotes

[1] This research was supported by grant NE-G-00-3-0016 from the National Institute of Education, Department of Health, Education, and Welfare.

A FURTHER STUDY OF THE EDUCATIONAL VALIDITY OF LEARNING POTENTIAL ASSESSMENT: RESPONSE TO OBJECTIVES-BASED CURRICULA

Milton Budoff

1976

Research Institute for Educational Problems
29 Ware Street
Cambridge, Massachusetts 02138

A further study of the educational validity of Learning Potential Assessment: Response to objectives-based curricula.

Purpose: This study sought to determine the validity of a group administered Learning Potential assessment with scores on four curricula which were based on behavioral objectives serving as the criterion measures. They were employed to guage the relative predictive power of the Raven Learning Potential Test, a nonverbal reasoning task, and IQ scores. The four units taught tasks in the perceptual, motor, academic, vocational-social and self-help areas.

The Problem of the Validity Criterion for Learning Potential Assessment

Traditional academic criteria are clearly inadequate because these children do so poorly academically.

In order to validate the results of the Learning Potential measures, one must develop data relevant to these children's abilities which are independent of their low scholastic attainments or aptitudes. But one of the serious difficulties in test validation is the problem of the appropriate, school-independent criterion, especially when the target of the efforts are children from non-Western, or non-middle-class backgrounds, who also may not speak the standard English of the middle class school. Sarason and Gladwyn (1958) state that one source of failure in the attempts to develop culture-fair tests, e.g., the Davis-Eells Games, was the lack of criteria which do not relate to school attainment but permit measurement of evidence of intelligent functioning.

Budoff, Meskin and Harrison (1971) developed a model for testing the educational significance of Learning Potential scores. They adapted an educational unit, developed by the Elementary Science Study (1966), that taught simple electrical concepts by allowing the student to manipulate flashlight batteries, bulbs, and copper wires. The investigators formulated a minimally verbal test by which to evaluate how much the students knew about electrical circuits prior to the course, and then retested them following teaching the unit. On one section of this test, the students could demonstrate their understanding by pointing to electrical circuits mounted on pegboards to indicate which bulb would light. In a second section, hey merely had to indicate whether the bulb specified in the circuit on the pegboard would light or not, and if so, whether it would be brighter, dimmer, or of the same brilliance as the standard unit of electricity (a lit one-bulb/one-battery circuit they held in their hands). This format was repeated in a written version in which students answered simple questions on printed diagrams that closely resembled what actual circuits looked like, and, in a second section, which utilized schematic or symbolic representations of these circuits as they might appear in a physics text.

The evaluation of the laboratory science classes supported the argument that special class students with high Learning Potential status on the Kohs block design task are educationally rather than mentally retarded. That is, the level of scores on the evaluation instrument did differentiate between the special and regular class students prior to teaching. After teaching, special and regular class children could be differentiated only on the basis of their Learning Potential status, not their class placement. This was most clearly evident when the performance of the low achieving CA controls was compared with the special class sample. The superior level of performance by the low achieving CA controls on the pretest was no longer evident on the posttest. There were no differences between the two groups when posttest or improvement scores were compared. Almost all the significant $\underline{F}$-ratios were lodged in the main effect for Learning Potential. Even more consistently than with the larger sample of CA controls, which included some high and average achieving students, the low achieving students' level of knowledge on the electricity test was indistinguishable from that of the special class students taken as a group, following exposure to a suitably designed curriculum unit.

A similar strategy of instruction and evaluation will be employed in this study except that the curricula describes skills and understandings in a rough sequential form, relevant to competencies in the perceptual motor, academic, vocational-social, and self-help areas. Each skill or understanding is phrased as a behavioral objective to be attained. Teaching procedures are suggested, which maximize the various modes of learning. These alternative teaching modes enable the child to utilize an experientially-based model for learning and thereby avoid the verbal-symbolic translation of ideas that these children have difficulty processing. This approach to learning enables the children to have a multi-sensory experience, e.g., a visual, aural-oral, tactile, learning situation. Rather than having to concentrate solely on auditory and visual symbol coding and decoding for learning, the child is able to make use of additional tactile experience, coupled with a "real world" model.

When children who are unable to cope with the verbal and structural demands of school and who consider themselves to be generally unable, are presented with manipulable materials, they perceive them in terms of their own past competencies with similar materials, e.g., toys, and are less likely to anticipate failure. The child who is apprehensive about failure becomes involved in active learning before he can remember to remain passive or aloof, and over succeeding days, should become progressively more receptive to other active and traditional learning activities.

We believe that children who have verbal deficits which impede their school performance can learn if the information and strategies are made compatible with their abilities. Such children "fail" in school, often beginning in kindergarten or first grade. They suffer the pains and frustrations of being unable to cope with the demands expected of them. They develop defensive postures designed to reduce the threat of incompetence, thus withdrawing further from the norm and confirming initial school evaluations. A child who is unable to read is destined to fail virtually every academic subject matter area he will encounter in elementary school. The proposed curriculum units use alternative teaching strategies and materials that minimize reliance on the child's reading and language ability; they allow the "retarded child" to begin on a sufficiently simple level so that he can experience success.

It is hypothesized that posttraining Raven scores will be most closely related to postteaching scores on the behaviorally-based units. That is, posttraining Raven scores will predict posttraining scores on units which make minimal verbal demands within empirical learning formats. This would be applicable to the perceptual motor and vocational-social objectives-based units and would constitute the most direct evidence of the validity of learning potential measurement. When working with academically oriented objectives, our hypothesis is that IQ will most accurately predict outcomes. Since the students in this study will have IQs higher than 45, it is likely that the learning related to the self-help objectives-based unit will not be predicted by either measure. It is our expectation that these children's self-help skills will not vary considerably on pretest, and be close to an asymptotic level for this unit. Hence, there will be little or no variability in the sample.

Method

Subjects

Data were collected in a large urban school system (New York City) between December and June of 1975. All subjects were special class students psychometrically classified as "educable mentally retarded." Raven's Progressive Matrices learning potential assessment was administered to 181 subjects (100 male, 81 female) before they participated in the behaviorally based curriculum that was to be initiated in their schools. Subjects ranged in age from 7 years 4 months to 16 years 10 months. (Mean = 11-10 years; SD = 2 years 2 months). Thirty-eight percent of the students were black; 70% of the subjects were in essentially self-contained classrooms while 30% were integrated into the regular school program for at least one-fourth of the school day.

Instruments

1. Learning Potential Assessment.

Raven Progressive Matrices, Sets A, A_B, B were group administered prior to and following the training procedures. In addition, Raven Sets C and D were given to all subjects nine years of age or older. A total of 11 school personnel--10 resource teachers and one special class teacher, conducted these testing and training sessions after thorough orientation to the materials and procedure by the RIEP staff. This training program covered four days. Following this, RIEP staff made two visits to each of the trainers' schools to observe their testing and training methods with the groups of children. The resource teachers are curriculum enrichment consultants to the special class teachers but also work with the students in these classes on a regularly scheduled basis.

Pre- and posttraining sessions were generally conducted with class size groups (N = 12-15) while training groups had an average of 5 students per group. The training procedures, which were completed in two 45-minute sessions, explained and demonstrated principles appropriate to pattern completion, orientation of complex elements in a design, and double classification problems. A training booklet which contained nontest problems of these types was distributed to each child. The trainers presented problems on slides from a Kodal Carousel projector on a blackboard surface. The requirements of each problem type were presented initially in meaningful designs, e.g., an American flag with a piece missing, and then geometric forms to attune the child to the basic format of the Matrices test problems. Individual children were called on to indicate the correct choice for each problem. A slide with the correct answer included allowed children to compare their choice with the correct one and if necessary, to correct it.

Subjects were assigned to training groups on the basis of pretest performance; students scoring at roughly similar levels were grouped together.

2. WISC or Binet IQ.

IQ information was obtained from school records. One hundred thirty four subjects had WISC scores while 35 had been tested on the Stanford Binet. For 12 students, the type of IQ test was not listed.

3. The Curriculum

The curriculum consists of instructional objectives written for four levels within the special education system; Trainable Mentally Retarded; Early Childhood; Elementary; and

Junior High. Subjects included in the present study, however, comprised only the latter two levels.

At each level, the curriculum is developed in four modular areas: Vocational-Social; Self-Help; Perceptual Motor; and Academic. Within these four modules, specific milestones have been defined. The milestones are further refined with respect to instructional objectives. Included for each objective is an instructional method and type of instructional resource, if necessary. The highest milestone in each of the modules represents more independent living or attainments. For example, in the Academic module of the Elementary curriculum, the first milestone which is number recognition has three objectives: naming written numbers 0-25; reproducing written numbers 0-25; and writing numbers 0-25. By milestone 13, 8 objectives are given for teaching concepts about money. These objectives include identifying coins and bills up to one dollar, making combinations of coins for a given value, writing money values using appropriate dollar and cent signs, and stating five functions of a bank.

The special class teachers assessed each student on each of the objectives in December and June. On each objective the teacher recorded whether the student had shown the behavior necessary to meet each objective.

Total number of objectives in each module varied across levels. For Junior High students there are 26 milestones with a total of 70 objectives within the Vocational-Social module; 12 milestones with a total of 61 objectives within the Self-Help module; and 25 milestones within the Academic module with a total of 136 objectives. At the elementary level the Vocational-Social module contains 4 milestones with 44 objectives, the Self-Help module comprises 9 milestones with 90 objectives and the Academic module is composed of 35 milestones with 160 objectives.

Conceptually, this type of program allows children to be programmed at the levels at which their current competencies dictate, with the best series of behavioral objectives very approximately ordered in the sequences described in the manual to the teachers. Conceptually, this type of program, then, allows for considerable individualization of programming by the teacher, and the expectation that children in the classes will be matched with children at similar levels of competence, if not programmed individually. Further, the program allows children to progress at their own rate, governed only by their need to satisfy the requirements of the particular behavioral objectives.

The extent to which the curricula were taught with these principles governing the teachers' programming for each child varied considerably, as did the number of behavioral objectives that each child was able to work on at any given time, the period of time allotted each child per day for this effort, and the extent to which children, when they completed the requirements of particular objectives, were immediately programmed into new ones. That is, the extent to which the children were truly individually programmed, and were allowed to move through the objectives at the rate at which they mastered them, varied considerably from classroom to classroom.

Results

One hundred and eighty-one subjects were included in the original sample. Of the subjects in this sample, complete data were not available for 59 cases. To insure the stability of the correlation coefficients and other statistics, results were obtained using the 121 subjects for which complete data were available.

Multiple regression techniques were used to analyze the data. Multiple regression analysis is a technique by which the unique and common contributions of various independent variables to the variation of some dependent variable can be determined. This technique permits a simultaneous assessment of the contributions of many independent variables to the variation in a dependent or criterion variable. The independent or predictor variables used in each of the following analyses were sex, race, age, posttraining Raven's Progressive Matrices, Binet IQ, Level of Special Class (Elementary, Junior High School), and degree to which student was in a self-contained or integrated program (The levels of the integration factor were: Self contained, 25%, 25-50%, and 50-75% integrated). The first six variables were used as covariates. The effects of containment will be considered once the effects of these six variables have been removed. The dependent, or criterion variables were the difference scores obtained by the subjects on the Raven Progressive Matrices (Sets A, A_B,B), Vocational-Social, and Academic tasks. These difference scores were obtained by subtracting the pretest score from the score that the student obtained following the period of teaching on the posttest. There were little or no differences among the number of objectives attained by these elementary and high school-aged EMRs on the self-help and perceptual-motor skills units following the teaching period, in large part because these students had largely attained many of the objectives prior to the onset of teaching.

The first dependent variable to be considered is the difference score obtained prior to and following training attuned to the Raven Progressive Matrices (Sets A, A_B,B). The results of the regression analysis is shown in Table 1. The only variables which significantly affected this difference score were level of posttraining Raven's score (R2), and the rough chronological age of the child as reflected in the level of his class (Junior High School or Elementary). Students with higher posttraining (R2) scores tended to show larger difference scores on this test. Students at higher levels also tended to exhibit larger "gains" or difference scores. There was no effect of the degree of integration factor influencing learning potential scores.

Table 1

A Summary of the Multiple Regression Analysis of Difference Scores Prior to, and Following Training on the Raven Progressive Matrices, Sets A, A_B, B

	Percent of Variance Explained		F-Ratio	Significance
Variable	Unique	Cummulative		
Sex	.000	.000	.086	$p > .1$
Race	.003	.003	.031	$p > .1$
Age	.001	.004	2.440	$p > .1$
Posttraining (R2)	.095	.099	7.502	$p < .01$
Binet IQ	.000	.099	.013	$p > .1$
Level of Class	.021	.120	2.919	$p < .1$
Degree of Mainstreaming	.002	.122	.323	$p > .1$

The second dependent variable considered was the difference scores obtained from the Vocational-Social test. The regression of the multiple regression analysis is shown in Table 2. These results indicate that there were three variables which affected the difference score. These variables were the subject's age ($F = 4.924$, $p < .05$), the posttraining (R2) score ($F = 7.33$, $p < .01$), and the type of integration ($F = 29.76$, $p < .001$). The older the student the lower the student's difference score tended to be. The relationship between the difference score and age was negative. The higher a student's posttraining (R2) score, the larger the difference score. The students who spend part of their school day outside the special class also showed larger gains in behavioral objectives. But it is likely that these are the more competent. The means for the four levels of integration mentioned previously, after adjusting for covariates, were 10.34, 15.87, 21.39, and 26.91, respectively. This effect exhibited multivariate significance.

Table 2

A Summary of the Multiple Regression Analysis of the Difference Scores on the Vocational-Social Objectives Curriculum

	Percent of Variance Explained		F-Ratio	Significance
Variable	Unique	Cummulative		
Sex	.007	.007	1.63	$p>>.1$
Race	.017	.024	1.72	$p>.1$
Age	.021	.045	4.92	$p<.05$
Posttraining (R2)	.069	.114	7.33	$p<.01$
Binet IQ	.004	.119	.75	$p>.1$
Level of Class	.000	.119	.43	$p>.1$
Degree of Mainstreaming	.182	.301	29.76	$p<.001$

The third dependent variable considered was the difference score in attainments on the Academic Objectives. The results obtained using this variable are shown in Table 3. Two variables contributed unique variance to a significant degree to the increase in number of academic objectives following teaching. These variables were posttraining (R2) score, and containment. The higher the posttraining (R2) score, the larger the increase in the number of academic objectives attained following teaching. Degree of integration also affected the gain score attained. Subjects in each of the four types of programs adjusted for covariates were 9.81, 12.05, 23.29, and 34.33, respectively for the continuum from special class, 100%, to largely integrated, which showed multivariate significance. Again, the more time the child spent in integrated placements, the greater the number of academic objectives he successfully completed during the six month period of teaching.

Table 3

A Summary of the Multiple Regression Analysis of ACAD Difference Score

	Percent of Variance Explained		F-Ratio	Significance
Variable	Unique	Cummulative		
Sex	.011	.011	1.843	*p* >.1
Race	.062	.074	.069	*p* >.1
Age	.022	.096	4.947	*p* <.05
Posttraining (R2)	.051	.147	2.845	*p* <.10
Binet IQ	.004	.151	.720	*p* >.1
Level of Class	.001	.152	.512	*p* >.1
Degree of Mainstreaming	.228	.381	42.022	*p* <.001

Discussion

The results following the educational training in objectives-based units showed three basic findings:

1. Response to the learning potential assessment procedure forecast the degree to which the EMR students mastered behavioral objectives in the vocational-social and academic objectives during the teaching period, without prejudice as to race, sex, chronological age, Binet IQ. This supports other findings suggesting that learning potential assessment may be a more culture-fair procedure than traditional IQ testing procedures. What is important to remember is that the students were uniformly low income students living in New York City, 38% of whom were black, the remainder, Caucasian.

2. The objectives-based units seemed to evoke considerable learning from the students regardless of their race or sex;

3. The children who spend more time outside the special class learned more objectives than those confined to the classes for almost all of their school day. It is likely, however, that these children are also more competent in ability, hence, the opportunities accorded them outside the special class.

As expected from prior work, the children's performance on nonschool (academic) dependent competencies was not predicted by the child's Binet IQ score, but was predicted by the child's response to the learning potential assessment. Of considerable interest is that the child's performance on the academically-based objectives was also predicted by his response to the learning potential assessment but not by his Binet IQ.

In neither of these curricula was response affected by sex or race of the child. This provides some support for the desirability of an objectives-based curriculum. It does seem to adequately pinpoint a child's competencies and learning needs and allow for skill-specific teaching to the identified needs. Children benefitted from the curricula areas regardless of race and sex.

For both areas of program, over the short span of this teaching effort, children who tended to spend some time out of the special class, and hence were probably more competent overall, also tended to master more of the behavioral objectives.

In summary, learning potential assessment predicted the EMR child's response to a nontraditional program of teaching even with academically relatdd objectives. Children who demonstrated increased proficiency following problem appropriate training to the Raven Progressive Matrices also showed increased response to the objectives-based units, when they represented a range of difficulty for the chronological and mental age range of the subjects involved. For the elementary and high school-aged pupils in this sample of EMR students, perceptual-motor and self-help skills were available to most students, and showed no variability among the members of the sample. It is likely that younger and/or more handicapped students would find these units of a sufficient range of difficulty to demonstrate a broad range of responsiveness to the objectives of these units.

Also: Reprinted in Journal of Educational and Psychological Measurement, 1974, 34, 977-981.

STUDIES IN LEARNING POTENTIAL

FACTOR STRUCTURES OF SPANISH-SPEAKING AND NON-SPANISH-SPEAKING CHILDREN ON RAVEN'S PROGRESSIVE MATRICES

by

Louise Corman and Milton Budoff

Volume 4, Number 73

1974

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

FACTOR STRUCTURES OF SPANISH-SPEAKING AND NON-SPANISH-SPEAKING CHILDREN ON RAVEN'S PROGRESSIVE MATRICES

Abstract

Item responses of Spanish- and English-speaking children on Raven's Coloured Progressive Matrices were submitted to principal components analyses and varimax rotations. Four factors were obtained which corresponded to readily identifiable problem types. The finding that three of the four factors were comparable for the two groups of subjects indicated the factorial invariance of the Raven test with children from different cultural backgrounds.

FACTOR STRUCTURES OF SPANISH-SPEAKING AND NON-SPANISH-SPEAKING CHILDREN ON RAVEN'S PROGRESSIVE MATRICES

Louise Corman and Milton Budoff

Research Institute for Educational Problems
Cambridge, Massachusetts

During recent years the Raven Progressive Matrices has become widely used as a measure of mental development of children and adults. Described by Raven (1965) as a test of "observation and clear thinking," the Coloured Progressive Matrices (Sets A, AB, B) were designed to assess the chief cognitive processes of children prior to the stage in which the intellectual capacity to reason by analogy is used as a consistent mode of inferential thinking. Raven considered the ability required to solve the puzzles in the Coloured Matrices to be independent of acquired knowledge or previously developed verbal skills.

In a previous study, Corman and Budoff (in press) factor analyzed item responses of retarded and nonretarded children and identified four types of problems represented by the 36 items in the Coloured Progressive Matrices: I--Continuity and Reconstruction of Simple and Complex Structures; II--Discrete Pattern Completions; III--Reasoning by Analogy; and IV--Simple Continuous Pattern Completions. This factor structure, which was replicated with two independent samples, was obtained for both retarded and nonretarded children and closely resembled item categories postulated by Raven (1965). From these results it was concluded that the Coloured Progressive Matrices is factorially invariant with children of different intellectual levels.

The purpose of the present study was to determine whether the pattern of item responses of Spanish-speaking students would result in a factor structure similar to that obtained with a sample of English-speaking subjects of normal intelligence. Comparability of factor structures of these two groups of subjects would provide evidence of the factorial invariance of this test with children from different cultural backgrounds.

Method

Subjects

The Spanish-speaking sample consisted of 228 students in three low-income urban communities in Massachusetts. The students were approximately evenly divided by sex and ranged in age from 6 to 17 years, with a mean age of 11-2 years (±2). Approximately 80 per cent of the students had been born in Puerto Rico.

The sample of 243 English-speaking subjects was drawn from low-income areas of a large urban community in Massachusetts. They ranged in age from 7 to 12 and were in the second through sixth grade. Approximately 60 per cent were male and about a fourth were black.

Procedure

Raven's Coloured Progressive Matrices (1956) were administered to all subjects in class size groups. Responses of the 36 items were dichotomized into correct and incorrect responses (1 = incorrect, 2 = correct). These 36 item scores were then submitted to a principal components analysis with 1.00 in the diagonal, followed by a varimax rotation of the four factors obtained. This procedure was carried out separately with item scores of Spanish-speaking and English-speaking subjects.

Results

For both groups of subjects, the four-factor solution was found to best satisfy the criteria recommended by Kaiser, presented in Harman (1967) [latent roots>1.00], and by Cliff and Hamburger (1967) [significant breaks in latent roots]. The four factors accounted for 48.4 per cent of variance for Spanish-speaking subjects and 42.5 per cent for English-speaking subjects; however, the largest amount of variance was contributed by Factor I for English-speaking subjects

(21.5 per cent) and Factor III for Spanish-speaking subjects (29.2 per cent). Factor loadings for both groups of subjects are presented in Table 1.

Insert Table 1 about here

Inspection of items with high loadings for both samples, as well as Burt's coefficients (Harman, 1967) which indicate the degree of comparability between the samples on each factor, revealed that item loadings on Factors I, II, and III were comparable for both groups of subjects. Factor IV, which contained items representing Simple Continuous Pattern Completions for English-speaking subjects, was not replicated by the Spanish-speaking subjects. The pattern of factor loadings indicated that many items which loaded highly on Factor IV for Spanish-speaking subjects were contained in Factor I for the English-speaking sample, and many items with high loadings on Factor IV for the English-speaking sample had high loadings on Factor II for the Spanish-speaking group. Items on both Factors II (Discrete Pattern Completion) and IV (Simple Continuous Pattern Completion) involved pattern completions rather than reorganization of the problem elements to attain a correct solution. Item difficulties indicated that large proportions of the Spanish-speaking students passed items which had high loadings on these two factors.

Discussion

Factor structures of Spanish- and English-speaking subjects on Raven's Coloured Progressive Matrices were found to be generally comparable. Three of the four factors contained many of the same items for both groups: I--Continuity and Reconstruction of Simple and Complex Structures, II--Discrete Pattern Completions, and III--Reasoning by Analogy. The fourth factor (Simple Continuous Pattern Completions) for English-speaking subjects was found to merge with Factor II for Spanish-speaking subjects. Items in Factors II and IV are similar in two

TABLE 1

Factor Loadings above .30 on Varimax Rotation for Two Groups of Subjects

	Factor I^a		Factor II		Factor III		Factor IV	
	Eng.	Span.	Eng.	Span.	Eng.	Span.	Eng.	Span.
Item	I^b	III	II	II	III	I	IV	IV
A1^c							-.84	
A2^c							-.75	
A3				.56			-.59	
A4			.39	.77			-.47	
A5				.72			-.64	
A6				.70			-.45	-.32
A7	.57	.38						-.41
A8	.39			.42				-.45
A9	.54			.33		.32		-.46
A10	.51			.30				-.55
A11	.48					.44		-.38
A12	.43				.30	.33		
AB1			.81	.70				
AB2			.74	.77				
AB3			.77	.72				
AB4	.47	.57						-.32
AB5	.63	.39						-.48
AB6	.68	.58						-.43
AB7	.58	.58	.31					-.32
AB8	.46	.52			.35	.35		
AB9	.47					.37		-.41
AB10	.42					.49		
AB11	.37	.58						
AB12					.59	.56		-.31

TABLE 1 (continued)

	Factor I[a]		Factor II		Factor III		Factor IV	
	Eng.	Span.	Eng.	Span.	Eng.	Span.	Eng.	Span.
Item	I[b]	III	II	II	III	I	IV	IV
B1		.52	.57	.48			-.32	
B2	.34	.49	.54	.57				
B3	.44	.57		.32				
B4	.60	.70						
B5	.63	.69						
B6	.43	.45				.46		
B7	.34				.33	.47		
B8					.72	.76		
B9					.82	.70		
B10					.76	.73		
B11					.80	.72		
B12					.63	.42		
% variance	21.5	4.7	11.0	10.2	5.7	29.2	4.3	4.3
Rotated SS	5.4	4.6	3.1	4.9	3.9	4.4	3.0	2.5
Burt's Φ_{pq}	.61		.71		.89		-.27	

[a] I--Continuity and Reconstruction of Simple and Complex Structures, II--Discrete Pattern Completion, III--Reasoning by Analogy, IV--Simple Continuous Pattern Completions.

[b] Actual factor.

[c] Every Spanish-speaking student passed this item.

respects: they are easy and involve pattern completions. Since ability to solve the Raven problems has been found to increase with age (Budoff, & Corman, 1973), it is likely that the older age range of the Spanish-speaking students in the sample accounted for their lack of difficulty with problems constituting these two factors. The overall similarity between factor structures of Spanish- and English-speaking children on this test provides evidence of the factorial invariance of Raven's Coloured Progressive Matrices with children from different cultural backgrounds. These results lend support to Raven's belief that ability to solve the puzzles on this test is independent of previously developed verbal skills.

References

Budoff, M., & Corman, L. The effectiveness of a group training procedure on the Raven Learning potential measure with children from diverse racial and socioeconomic backgrounds. *Series in Learning Potential*, 1973, *3*, 58.

Cliff, N., & Hamburger, C. D. The study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin*, 1967, *68*, 430-445.

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven's Progressive Matrices. *Educational and Psychological Measurement*, in press.

Harman, H. H. *Modern factor analysis*. Chicago: The University of Chicago Press, 1967.

Raven, J. C. *Coloured progressive matrices*: Sets A, AB, B. London: H. K. Lewis & Co., Ltd., 1956.

Raven, J. C. *Guide to using Raven Progressive Matrices*. London: H. K. Lewis & Co., Ltd., 1965.

Footnote

[1] This research was supported by Grant OEG-1-72-0020 (509), Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, Office of Education, Grant NE-G-00-3-0016 from the National Institute of Education, and Grant R01 MH 18553 from the National Institute of Health, all under the Department of Health, Education, and Welfare.

PLEASE REFER TO ED 108 436
Also: Reprinted in Interamerican Journal of Psychology, 1976, 10, 13-24.

STUDIES IN LEARNING POTENTIAL

AN EDUCATIONAL TEST OF LEARNING POTENTIAL ASSESSMENT WITH SPANISH-SPEAKING YOUTH

by

Milton Budoff, Louise Corman, and Alex Gimon

Volume 4, Number 71

1974

Research Institute for Educational Problems, Inc.
29 Ware Street, Cambridge, Mass. 02138

PLEASE REFER TO ED 108 436

AN EDUCATIONAL TEST OF LEARNING POTENTIAL ASSESSMENT WITH SPANISH-SPEAKING YOUTH

Abstract

The learning potential (LP) procedure represents an alternative method of measuring the general ability of Spanish-speaking students who tend to score low on traditional IQ tests. Postteaching scores on an electricity curriculum unit test were used as criteria to compare the relative predictive power of LP and IQ measures for Spanish-speaking students. Subjects were administered the Raven LP procedure, the Semantic Test of Intelligence, the WISC Performance Scale in Spanish, and the WISC Vocabulary Subtest in Spanish and English. Before and after participation in an electricity unit, subjects took the electricity unit evaluation instrument. Posttraining Raven LP scores significantly predicted performance on the minimally verbal symbolic level of the electricity instrument. IQ scores were not positively related to postteaching electricity scores. The LP procedure resulted in increased levels of performance on a reasoning task.

PLEASE REFER TO ED 108 436

AN EDUCATIONAL TEST OF LEARNING POTENTIAL ASSESSMENT WITH SPANISH-SPEAKING YOUTH [1]

Milton Budoff, Louise Corman, and Alex Gimon

Research Institute for Educational Problems

Budoff, Gimon, and Corman (1974) have set forth a rationale for an alternative method to the IQ test for measuring intelligence of Spanish-speaking children, where this construct is defined as the ability to learn and profit from appropriate experience. Learning potential (LP) assessment utilizes a three-stage procedure which includes pre-and posttraining administrations of a nonverbal problem solving task with training relevant to the tasks interpolated. The entire procedure may be group or individually administered. The strategy is based on the premise that low income and/or minority group children differ in familiarity and experience with particular tasks, have a negative expectancy of success while taking tests, and are less effective in spontaneously developing strategies appropriate to solving the often strange problems on a test. Training helps these children develop a sense of competence on the task by providing them with problem-relevant strategies in a context of positive support.

Posttraining scores are indicative of general ability to reason. The posttraining score is hypothesized to relate to students' performance on tasks or curricula which permit them to operate in their areas of conceptual or cognitive strength, as opposed to areas of weakness, such as the verbal-conceptual domain predicted by IQ scores. The purpose of the present study was to test this hypothesis with Spanish-speaking children.

Previous research with IQ-defined EMRs (Budoff & Corman, in press) has shown that pretraining learning potential scores were related to verbally-biased IQ scores and to social and demographic factors associated with high academic risk. Posttraining scores were found to be related to other performance ability measures, e.g., WISC performance IQ, but were not related to academic risk factors.

Similar findings have been obtained with Spanish-speaking children. Budoff, et al. (1974) administered learning potential and IQ measures to 188 Spanish-speaking children in grades 1 to 6. The WISC administered in Spanish yielded low estimates of intellectual ability (mean performance IQ score = 86 ± 15; verbal IQ, 77 ± 17), and depicted the students as slow learners or mentally retarded. The pretraining learning potential score on Raven's Progressive Matrices (1958) was correlated significantly with the WISC IQ scores and averaged at the 25th percentile for the chronological age of these subjects. By contrast, the average posttraining score was at the 67th percentile. The correlation of posttraining scores with the WISC verbal score was not significant, while there continued to be a strong relationship to WISC performance IQ.

Budoff, et al. (1974) used a standardized achievement test, the Inter-American Test of General Abilities (IAGAS), as a criterion measure to compare the relative predictive power of IQ and LP scores. The IAGAS measures verbal, nonverbal, and numerical achievement in Spanish and English. Posttraining LP scores, as well as PIQ, were significantly related to nonverbal and numeric achievement scores in both Spanish and English, while WISC verbal IQs predicted only verbal achievement in the same language. Hausman (1972) also reported a significant relationship between LP scores and achievement with samples of Mexican-American children. Interpretation of these findings, however, must take into account that any academic achievement criterion of validity is inadequate for children who perform so poorly in Spanish and English achievements. The mean achievement level of the sample in the Budoff, et. al (1974) study was below normative expectation in Spanish and English. To compare the relative power of IQ and LP measurement, the criterion should permit children to display their reasoning capability in situations not adversely affected by poor scholastic attainments.

The present study employs this kind of criterion. Budoff, Meskin, and Harrison (1971) developed a model by which to test the relative power of IQ and LP assessment with IQ-defined educable mental retardates on educationally relevant curriculum units. They reasoned that IQ-defined educable mental retardates who profited markedly from the learning potential training would also learn when the curriculum permitted maximal reliance on their nonverbal reasoning ability and a minimal demand on their poor verbal-expressive competence. To test the educational significance of learning potential scores with educable mentally retarded and average-IQ, low income students, the students were taught concepts of electricity by manipulating flashlight batteries, bulbs, and copper wires. The investigators formulated an evaluation instrument that required a minimally verbal response (pointing or one word answers) in a multiple choice format and determined the students' acquisition of knowledge in a pre-to-postteaching design. The students were also asked to give reasons for their choices for a subsample of items. The investigators hypothesized that these students' ability to provide reasons for their choice would be less adequate than their empirical understanding of the problems.

Results of the Budoff, et al. (1971) study indicated that high able EMR students, by the learning potential criterion, learned more from the electricity unit than students who had not improved their LP scores following training on the learning potential procedure. Performance on the electricity evaluation instrument differentiated between the special and regular class students prior to teaching. Following teaching, special and regular class children could be differentiated only on the basis of their learning potential status, not their class placement. However, special class students were significantly poorer than their regular class agemates in explaining the reasons for their correct minimally verbal responses. This difference in performance was consonant with the low IQ scores of special class students.

The present study employs the same model to test the educational significance of learning potential assessment with Spanish-speaking students. The electricity unit was translated and taught in Spanish and the evaluation test was administered in Spanish prior to and following the course. This instrument allowed comparisons at two levels of abstraction within a minimally and maximally verbal mode of response. The minimally verbal sections of the test required responses to actual electrical setups mounted on pegboards (concrete level) and to electrical problems presented as two-dimensional drawings (representational or symbolic level). In the verbal sections of the test the students were asked to explain in Spanish the reasons for their choices to a subset of pegboard items. Responses were scored for verbal-descriptive and verbal-conceptual competence.

This study hypothesized that postteaching scores on the minimally verbal section of the electricity test would be predicted by posttraining learning potential scores but not by IQ scores. Verbal explanation scores on the electricity instrument following teaching were expected to be more highly related to scores on a language-related training-based measure than to WISC vocabulary scores.

Method

Subjects

Data were collected in a transitional bilingual school in an urban community in New England from March, 1972, to May, 1973. Fifty-four Spanish-speaking students in classes equivalent to second through sixth grades were administered the WISC IQ test, the Semantic Test of Intelligance, and Raven's Progressive Matrices learning potential assessment in Spanish. After these tests had been given, students participated in the electricity classes as part of their general science requirement. The age range of this group was 6-7 to 13-8 years, with a mean age of 9-10 (±1-9) years. The sample consisted of 32 boys and 22 girls. The mean WISC full scale IQ in Spanish (86 ± 18) indicated that many would be psychometrically classified as "slow learners" or "educable mentally retarded." The students came from low income homes, with the majority of the families receiving public assistance payments.

Spanish bilingual classes had been in existance for three years in this community. Most of the older students had English language instruction in the earlier grades. The younger children (second and third grades) had been in Spanish language classes from the start of formal schooling and were learning English as a second language.

Instruments

1. Learning Potential assessment. Raven Progressive Matrices, Sets A, AB, B (1956) were group administered in Spanish prior to and following the group administered training procedure offered in one 45-minute session. The training procedure explained and demonstrated principles appropriate to pattern completion,

orientation of complex elements in a design, and double classification problems, which are the most readily identifiable problem types in these problem sets. Corman and Budoff (in press) have confirmed that these problems are distinct types. A training booklet which contained nontest problems dealing with these types was distributed to each child. The trainer presented the problems on 2 X 2 slides from a Kodak Carousel projector on a blackboard. The students were required to draw in the missing element for the design before they looked at the six choices presented on the lower half of the page. This procedure was particularly useful in helping these children understand the double classification problems. During development of the procedure, it was found that children could easily derive one attribute at a time, but often did not hold the first attribute in mind while they derived the second relevant attribute. The child's understanding was facilitated by having him draw the relevant attributes, one at a time, as he derived them. This procedure helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems mentally with very little trouble.

The requirements of each problem type were presented in meaningful designs initially, e.g., an American flag with a piece missing, and then a geometric form to attune the child to the basic format of the Matrices test problems. Individual children were called to indicate the correct choice, and to give reasons for their choice. A slide with the answer included allowed them to compare their choice and to correct it, if necessary. A manual in Spanish translation with the instructions and procedures is available (Gimon, Budoff, & Corman, 1974).

2. The Semantic Test of Intelligence (STI) was group administered to all students with the standardized procedures developed by Rulon and Schweiker (1953).

Any necessary instructions were given in Spanish. The STI is a language-related measure which uses geometric symbols to represent a pictured object or action as an analogue to reading words in a sentence. It is administered as a timed test and consists of 217 items, including 109 items with one symbol, 49 with two symbols, 36 with three symbols, and 23 with four symbols. The symbols for each noun and verb are introduced as single symbols, and defined by accompanying pictures in a multiple choice format on tuition pages which are not scored. Tuition pages are used to introduce 2-, 3-, and 4-symbol "sentences." Instructions are pantomimed. The symbols and the pictures which define their meaning are presented on each double page so that memory for the meaning of the symbol is not a factor influencing performance.

The STI was originally developed as a measure of military trainability with illiterate recruits who had failed the literacy requirement for entry into the Marine Corps. Validity of this test as a language-related measure has been evidenced by Rulon and Schweiker's (1953) finding that recruits who did well on the STI also successfully completed a literacy course to meet the eligibility requirements of the Marine Corps. Proficiency on this training-based language-related task, more than WISC vocabulary scores was hypothesized to relate to the child's verbal competence in explaining the reasons for his empirical understanding of electricity.

3. WISC IQ. The students were individually tested on the WISC performance subtests in Spanish and the WISC vocabulary subtest in Spanish and English, Puerto Rican version (1954) and scores were prorated to provide an estimate of verbal IQ.

Electricity Unit

All students participated in the electricity course as adapted by Budoff, and Meskin, 1970. This unit teaches simple concepts of electricity by having the child experiment with his own equipment (flashlight batteries, bulbs, and various types of wire) and draw conclusions based on the results of his observations of these experiments. The unit employed earlier was translated and taught in Spanish. All workbook materials were in Spanish. The classes consisted of 10 to 12 students, grouped by grade level, and were conducted during two 45-minute sessions a week for a seven-week period. Electricity classes were organized following completion of the IQ and LP testing.

Children learned what a simple circuit is and what requirements are necessary for making a bulb light. They studied the results of varying amounts of voltage and amperage in an elementary way by making and observing different circuits with bulbs and batteries arranged in series or in parallel circuits. They noted the effects of resistance in wires of different materials, lengths, and thicknesses, and in different kinds of bulbs. Conductors and insulators, both solid and liquid, were observed. Lessons in diagramming with electrical symbols were included in the unit. Language was kept simple--the children learned few formal electrical terms, and formal explanations of concepts such as electrical resistance were minimized. No attempt was made to offer a comprehensive theory of electricity. A description of the course and the teacher's manual and student's workbook are presented in Budoff and Meskin (1970).

Electricity Evaluation Instrument

Before and after participation in the electricity unit, all students were given an instrument developed to evaluate their understanding of electricity (Budoff & Meskin, 1970). The test was designed to permit responses at a concrete and abstract level in minimally and maximally verbal response modes. The concrete level of the minimally verbal section of the test consisted of actual electrical setups mounted on masonite pegboards. This presentation of materials, made familiar by the course, increased the child's opportunity to demonstrate empirically his comprehension of the concepts of electricity. The symbolic level of the minimally verbal section presented problems as diagrams which looked like the setups (realistic), and as schematic diagrams, to test the child's ability to apply his acquired knowledge in more abstract presentations. The symbols presented in the schematic diagrams were learned in the course and were explained in the test booklet. The pegboards section was administered individually, and the diagrams section in groups.

The verbal section of the test required subjects to explain their answers to 16 selected questions on the pegboard section. These answers permitted an evaluation of the subject's ability to verbalize his behaviorally demonstrated understanding. Answers were scored for two levels of response: the verbal descriptive level, which represented the child's ability to describe the salient characteristics of the circuit, and the verbal conceptual level, which indicated the student's causal understanding of why a particular bulb would or would not light.

Four scores were obtained for each student on both the pre- and postteaching test: one for each level of response within the minimally and maximally verbal response modes. Scores on the pegboards (concrete level) and diagrams (symbolic level) were number correct. The total score possible on the pegboards was 60 and on the diagrams was 62. Each of the 16 verbal explanations was scored on a 4-point scale so that the total possible score for each verbal level was 64.

Results

The means and standard deviations for the WISC PIQs were in the average range (93.2 ± 16). The prorated means for the WISC verbal IQ was 81.5 (± 21) and 50.6 (± 18), for the Spanish and English administrations, respectively. The mean Semantic Test of Intelligence score was 155.2 (±41.9). The Raven scores increased with training, from 21.2 (± 7.6) to 25.1 (± 7.1)

The pre- and postest means for electricity test scores presented in Table 1 indicate the students did know more about the application of the principles of electricity after the course.

Insert Table 1 about here

Predictors of Postteaching Performance on Electricity Instrument

Four stepwise multiple regression analyses on postteaching scores on the electricity instrument were performed to test the relative predictive power of Learning Potential and IQ scores on each level of response. The independent variables were posttraining Raven score, pretest score on the electricity measure, WISC Performance IQ, WISC Spanish Vocabulary, WISC English Vocabulary, Semantic Test of Intelligence score, age, and sex. These four regression analyses were repeated with pretraining Raven scores substituted for posttraining Raven scores to compare the predictive power of these two scores on the criterion variables. In all regression analyses, the order of entry of independent variables into the equations was determined by the amount of variance accounted for by each variable.

Posttraining scores on the learning potential procedure significantly predicted postteaching diagram scores, the minimally verbal test of symbolic electrical knowledge (See Table 2). By contrast, the students' scores on three IQ measures, i.e., pretraining Raven, WISC Performance IQ and WISC Spanish vocabulary scores did not significantly predict any of the four postteaching

TABLE 1

Means and Standard Deviations on Electricity Measures

	Pretest		Posttest	
	$\bar{X}$	SD	$\bar{X}$	SD
Minimally verbal mode				
Pegboards (concrete)	30.98	7.02	49.56	3.09
Diagrams (symbolic)	17.74	6.28	44.93	8.22
Verbal mode				
Descriptive reasoning	15.67	5.53	40.17	5.92
Conceptual reasoning	8.65	4.09	28.48	7.18

electricity scores. This finding supported the hypothesis: after LP training had provided the experiential knowledge and skills which these low income, linguistically different students lacked, posttraining LP scores were better predictors of success on the curriculum unit than IQ scores.

Insert Table 2 about here

The Semantic Test of Intelligence score which was hypothesized to relate to the verbal explanation scores, significantly predicted postteaching scores only on the pegboard section (the minimally verbal concrete measure), and did not relate to either verbal score. The sign of the beta weight for STI in this equation was negative. It is likely that the STI functioned as a suppressor variable: it was highly positively correlated with WISC performance IQ but correlated zero with the posttest pegboard score.

English WISC vocabulary score was significantly and negatively related to verbal conceptual reasoning. That is, the lower the facility the students had in English vocabulary, the higher their conceptual reasoning in Spanish, as indicated by their ability to provide reasons for their empirical understanding of a problem.

Sex was not a significant predictor of any of the four criterion measures. Age was a negative predictor of minimally verbal symbolic, verbal descriptive, and verbal conceptual scores. Younger students scored higher on these measures than older students.

Improvement Following Electricity Instruction

To determine change in performance following participation in the electricity unit, scores on the two modes of response (minimally and maximally verbal) were submitted to separate repeated measures analyses of variance. In each analysis,

TABLE 2

T-Tests of Independent Variables in Multiple Regressions on Postteaching Electricity Scores

	Dependent variables							
	Minimally verbal mode				Verbal mode			
	Pegboards (concrete)		Diagrams (symbolic)		Descriptive reasoning		Conceptual reasoning	
Independent variables	Beta weight	t-value	Beta weight	t-value	Beta weight	t-value	Beta weight	t-value
Electricity pretest	.463	3.17**	.348	2.65*	.244	1.69	.086	0.69
Age	-.254	-1.66	.446	-3.47**	-.437	-2.72**	-.574	-4.20***
Posttraining Raven	.052	0.36	.391	2.93**	.138	0.84	.238	1.67
WISC performance IQ	.354	1.77	-.025	-0.14	-.025	-0.11	-.001	-0.00
WISC Span. vocab.	.264	1.48	-.038	-0.23	.104	0.51	-.107	-0.60
WISC English vocab.	-.064	-0.46	.016	0.12	.051	0.32	-.299	-2.16*
Sex	-.082	-0.56	.047	0.35	.045	0.57	-.075	-0.52
STI	-.329	-2.03*	.045	0.27	-.145	-0.77	.049	0.30
r^2	.377		.462		.190		.386	
F (*df* = 3/45)	3.40**		4.83***		1.32		3.54**	

*p <.05

**p <.01

***p <.001

the between-subjects factor was Sex, and the two within-subjects factors were Test Session (pre- and postteaching) and Test Level (i.e., concrete and symbolic for the minimally verbal mode and descriptive and conceptual for the verbal mode).

The results of these analyses are presented in Tables 3 and 4 for each response mode. The significant main effect for Test Session in both analyses showed highly significant gains from pre- to posttest for both modes of response ($F = 682.16$, $df = 1/52$, $p < .001$ for the minimally verbal mode; $F = 514.21$, $df = 1/52$, $p < .001$) for the verbal mode). Significant differences in levels of response competence within each mode were also obtained: across test sessions scores on the concrete level were higher than on the symbolic level, and descriptive scores were higher than conceptual scores (Table 1). The effects of Sex and interactions involving Sex were not significant in either analysis.

Insert Tables 3 and 4 about here

Significant Test Session X Test Level interactions were obtained for both modes. Lindquist's (1956) test of critical differences between means on the two minimally verbal levels showed that, while pretest scores on the concrete level were higher than on the symbolic level, scores on the two levels did not differ markedly after teaching. The means in Table 1 indicate that, following teaching, scores on the pegboards approached a ceiling and showed little variability. This finding indicated mastery of the concrete mode by most students after participation in the electricity course. A more marked increase in competence was evident on the symbolic mode. These results confirmed earlier findings of Budoff, et. al., (1971) which indicated a generalization of effect from the concrete pegboard format to the schematic diagrams and a significant consistency of response from one format to the other following teaching.

TABLE 3

Repeated Measures ANOVA on Minimally Verbal Electricity Scores

	SS	df	MS	F
Sex	28.45	1	28.45	0.46
Unit	3233.88	52	62.19	---
Test Level (concrete vs symbolic)	4343.48	1	4343.48	91.37*
Sex X Test Level	55.61	1	55.61	1.17
Test Level X Unit	2471.92	52	47.54	---
Test Session (pre-post)	27013.54	1	27013.54	682.16*
Sex X Test Session	21.78	1	21.78	0.55
Test Session X Unit	2059.19	52	39.60	---
Test Level X Test Session	1057.34	1	1057.34	63.89*
Sex X Test Level X Test Session	59.19	1	59.19	3.58
Test Level X Test Session X Unit	860.52	52	16.55	---
Total	41204.88	215	191.65	

*$p < .001$

TABLE 4

Repeated Measures ANOVA on Verbal Electricity Scores

	SS	df	MS	F
Sex	7.06	1	7.06	0.14
Unit	2650.48	52	50.97	---
Test Level (descriptive vs conceptual)	4466.03	1	4466.03	199.89*
Sex X Test Level	14.48	1	14.48	0.65
Test Level X Unit	1161.84	52	22.34	---
Test Session (pre-post)	25661.91	1	25661.91	514.21*
Sex X Test Session	0.42	1	0.42	0.01
Test Session X Unit	2595.08	52	49.90	---
Test Level X Test Session	304.44	1	304.44	24.15*
Sex X Test Level X Test Session	10.44	1	10.44	0.83
Test Level X Test Session X Unit	655.56	52	12.61	---
Total	37527.73	215	174.55	

*p <.001

Testing the differences between the means for the two verbal levels showed that the students were more competent in describing the characteristics of the circuits than they were in providing causal explanations for the electrical phenomena. Both the mean pre- and postteaching scores on description were higher than the corresponding means on conceptual expression. The significant Test Session X Test Level interaction indicated that the mean increase from pre- to posttest was greater on descriptive than on conceptual reasoning (Table 1); however, the fact that the students' ability to verbalize reasons for their choices at both levels increased markedly indicated that they had little difficulty at either level of performance in their native language.

Discussion

Optimizing the manner in which measures of general ability to profit from experience, i.e., intelligence, are administered, shows promise of providing a less biased indicator of general ability among Spanish-speaking students than that given by the traditional intelligence tests. The Spanish version of the WISC IQ test, given in the traditional, single administration format, revealed the usual finding that many of these students perform in the dull normal range. This finding was particularly evident for their verbal IQ even when administered in Spanish. Clearly these students are not competent when their spontaneous productions on school-oriented language or problem-solving tasks are used as the basis for estimates of intelligence. But these estimates reflect the difference in their language, as well as the difficulties low income students, regardless of linguistic or cultural background evidence when confronted with the middle class style of reasoning problems in tests, and presumably, in school.

By contrast, the learning potential procedure which shows these low income Spanish-speaking students how to perform more effectively on a reasoning task in a competence-inducing context, results in higher levels of performance on the reasoning task and predicts their ability to profit from the electricity curriculum. This finding provides additional direct educational validity for the efficacy of a training based assessment as embodied in the learning potential procedure. Budoff, et al., (1974) demonstrated that posttraining LP scores, as well as WISC Performance IQ scores presented in Spanish and English, did correlate with nonverbal and numeric achievement test scores. Verbal IQ predicted only verbal achievement in the same language. Achievement measures, however, tend to reflect the poor scholastic attainment of these students and are poor evidence of potential ability to learn under more suitable circumstances. The electricity unit, like the LP procedure, permitted these students to display their reasoning capability in situations not adversely affected by their prior poor achievements in school. After participation in the unit, students demonstrated marked improvement in both minimally verbal and verbal modes of response.

Validity of the LP procedure was indicated by the finding that posttraining LP scores were significantly related to performance on the minimally verbal symbolic level of the electricity evaluation instrument, whereas IQ scores were not positively related to any electricity scores attained following participation in the electricity unit, partially confirming the hypothesis. The finding that posttraining LP scores did not significantly predict postteaching scores on the concrete level of the electricity instrument reflects the ceiling and low variability of students on that section of the test. That is, after instruction, most students demonstrated mastery on the concrete level.

The expectation that postteaching verbal scores on the electricity instrument would be related to scores on the Semantic Test of Intelligence, a language-related training-based measure was not borne out by the results of this study. Students above the third grade level attained a ceiling on the test. While this test may possess validity with illiterate adults (Rulon & Schweiker, 1953) its usefulness with children spanning a broad age range would appear limited.

The finding that younger children attained higher postteaching scores on three of the four electricity measures is inconsistent with results of many studies, and may be attributed to the fact that all materials, discussion, and test administrations were conducted in Spanish. The older students in this sample had spent their primary grades in English-speaking classrooms. Also, they ranged in age to 15 years but were still attending a grade school. They may have been "turned off" to new learning situations, however stimulating. The younger students had been taught mainly in their native language since they started school. Their ability to express in Spanish an adequate verbal explanation for their nonverbal choice may reflect greater comfort and adequacy in their native language. This reason may also explain the negative relationship between English vocabulary scores and conceptual reasoning in Spanish; i.e., that relationship is probably stronger for the younger students.

Scores of this sample on a traditional IQ test indicated that these students' aptitude for academic tasks was low and that low school achievement could be expected. From these scores the inference is commonly made that they are less intelligent, i.e., they cannot perform more adequately. One major response to this clearly discriminatory judgment has been to restandardize the traditional IQ and achievement tests so they will normatively reflect the scores of low income and/or bilingual students (Mercer, 1972; Diana v. State Board of Education, 1970). But lowering the norms will only lower the expected level of performance, which will be reflected in a lower level of educational stimulation and challenges accorded these children. The result will be a de facto outcome that they are less able because their educational attainments are low. Lowering the raw score norms will reduce the incidence of mislabeling as mentally retarded, but will fail to close the gap between the child's acquired competencies and the deficiencies perceived by the school, and by extension, the society (Meyers, 1973).

A training-based assessment shows promise of providing a culture fair assessment of ability to profit from experience and meeting these objections. The criterion of adequate performance is maintained because the posttraining scores are evaluated in the context of the child's normative age group for the general population. But the recognition that low income and/or non-English-speaking children suffer considerable handicaps in responding to the test problems has resulted in a significant reorientation in the testing procedures. Training, i.e., helping the students to perform more effectively on the reasoning problems, enables the capable student to reorient his problem-solving abilities to the demands of middle class biased procedures. Further, embedding the test within this competence-enhancing learning situation helps the child who expects to fail cope with the anxiety generated by the test situation.

The educational implications of these findings should be clear. Many of these children are intellectually competent but the conditions must be structured for them in light of their prior cognitive history, and their present functioning styles. Especially for children who are marginally adequate achievers in classes conducted in a language and a culture foreign to them, curricula must be presented in the modalities that will tap these children's learning strengths. These curricula would tend to mitigate the failure set that low achievers tend to develop by enhancing their sense of competence in school. Manipulative, experiential learning units are often dismissed as irrelevant to the needs of problem learners because they do not foster improved literacy. However, they do engage these children in constructive learning. This provides a set of common experiences within a classroom that can be utilized for active language exchanges and for experienced-based reading materials. Talking about what they are learning by their own active efforts provides a basis for communications that should enhance these lower class students' ability to use language as an instrument of analysis and synthesis in problem solving. It is this relative failure that characterizes the major source of the verbal deficits among lower class students, when contrasted with children from middle class homes (Bruner, 1971).

This educational strategy is premised on the fact that their teachers will regard them as more able, and that they will structure programs appropriate to more adequate learners. The particular contribution of the training-based assessment score, LP, in contrast to the low IQs, is the demonstration that these students are often more intellectually competent than is generally assumed.

References

Bruner, Jerome S. *The relevance of education*. New York: W. W. Norton & Company, Inc., 1971.

Budoff, M., & Corman, L. Demographic and psychometric factors related to improved performance on the Kohs learning potential procedure. *American Journal of Mental Deficiency*, in press.

Budoff, M., Gimon, A., & Corman, L. Learning potential measurement with Spanish-speaking youth as an alternative to IQ tests: A first report. *Studies in Learning Potential*, *4*, 70, 1974.

Budoff, M., & Meskin, J. An educational test of the learning potential hypothesis with adolescent mentally retarded special class children. Final Report, 1970, Grant 32-31-0000-6019, Office of Education. Eric No. ED 033-484.

Budoff, M., Meskin, J, & Harrison, R. H. Educational Test of the learning potential hypothesis. *American Journal of Mental Deficiency*, 1971, *76*, 159-169.

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven's Progressive Matrices. *Journal of Educational and Psychological Measurement*, in press.

Diana v. State Board of Education, C-70 37 RFP, District Court for Northern California (February, 1970).

Gimon, A., Budoff, M., & Corman, L. Applicability of learning potential measurement with Spanish-speaking youth as an alternative to IQ. Final Report, 1974. Grant OEG-1-72-0020(509), Office of Education.

Hausman, R. M. Efficacy of three learning potential assessment procedures with Mexican-American educable mentally retarded children. Unpublished doctoral dissertation, George Peabody College for Teachers, 1972.

Lindquist, E. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1956.

Mercer, J. Sociological perspectives on mild mental retardation. Paper presented at the First Annual Study Conference in School Psychology, 1972.

Meyers, C. E. The school psychologist and mild retardation- Report of an ad hoc committee. **Mental Retardation**, 1973, 11(1), 15-20.

Raven, J. C. Coloured progressive matrices: Sets A, AB, B. London: H. K. Lewis & Co., Ltd., 1956.

Raven, J. C. Standard progressive matrices, Sets A, B, C, D, and E. London: H. K. Lewis & Co., Ltd., 1958.

Rulon, P., & Schweiker, R. Validation of a non-verbal test of military trainability. Final Report, June, 1953, Department of the Army.

Footnote

[1] This research was supported by Grant OEG-1-72-0020 (509), Grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, Grant NE-G-00-3-0016 from the National Institute of Education, and Grant RO1 MH 18553 from the National Institutes of Health. Special thanks are due to John Corcoran and Ann O'Donnell, Directors of the bilingual programs in Worcester and Lowell, Mass., for their cooperation. Thanks also to Dr. Robert H. Harrison for statistical consultation.

Also: Reprinted in American Journal of Mental Deficiency,
1976, 81, 49-57.

riep

PRINTS

Optimizing Test Performance of the Moderately and Severely Mentally Retarded

Milton Budoff and James L. Hamilton

RIEP-Print Number 85

May, 1975

RESEARCH INSTITUTE FOR EDUCATIONAL PROBLEMS, INC.
29 WARE STREET ▪ CAMBRIDGE, MASSACHUSETTS 02138 ▪ TELEPHONE (617) 868-0360-2

Abstract

This study provided further evidence for the validity of a learning potential assessment procedure with institutionalized moderately and severely retarded adolescents and adults. Significant positive correlations were obtained between psychometric and learning scores, attendant and teacher ratings of ability, and the posttraining scores on the modified Kohs Extended Learning Potential procedure. In addition, performance on this test-train-test procedure was compared with a train-within-test format for two different tasks: training embedded within the administration of the Leiter International Scale and a formboard version of Raven's Progressive Matrices. The students responded equally to the two formats. Stanford-Binet IQs were least predictive of performance on the three learning potential measures and were unrelated to teacher and attendant ratings of ability. The implications of these data are discussed with particular attention to the potential advantages of the train-within-test model.

Optimizing Test Performance of the Moderately and Severely Mentally Retarded

Although IQ tests such as the Stanford Binet (SB) and Peabody Picture Vocabulary Test (PPVT) have been widely used with the moderately and severely retarded for placement and other purposes (Silverstein, 1963), numerous investigators have questioned the validity of these tests with this population. Jensen (1963) has suggested that conventional IQ tests are, fundamentally, achievement tests which indicate what the child has learned outside the test situation and hence are of limited value for those children who have not had appropriate opportunities for learning those skills measured by the tests.

The usefulness of the standard individual intelligence tests with these persons has been questioned for several reasons: (a) the response requirements of many assessments penalize the moderately and severely retarded, a population replete with multiple handicapping conditions (O'Connor, Justice, & Payne, 1970); (b) the single administration format of standard tests untenably assumes that the person's initial responses are reliable (Barrett, 1965; Zeaman & House, 1963; Jensen, 1963), and that the person has a normal experiential background from which he has spontaneously learned knowledge and skills common to persons of the same age (Clarke & Clarke, 1967); and (c) data from conventional tests are frequently inadequate as predictors of learning ability (Gardner, 1945; Talkington, 1967).

What is required are new assessment formats which could be used to indicate the ability of moderately and severely retarded persons to learn and profit from instructional experiences. Because of the high incidence of speech and language problems among these persons, the assessment should be nonverbal. The expectancy of failure in evaluative situations dictates that the procedure be couched in a supportive, nonpenalizing context to maximize the person's efforts. To minimize the negative effects of the differences in these persons' experiences on their efforts, many investigators (Luria, 1961; Penrose, 1934; Tobias, 1960; Vygotsky, 1934/1962) have suggested the incorporation of instruction within the assessment procedure so that the person's familiarity with the task can be assured. With the exception of Schucman's (1960) study and Hamilton and Budoff (1974) studies, however, the inclusion of training within an assessment model has been neglected with the moderately and severely retarded.

Hamilton and Budoff (1974) applied the learning potential model advanced by Budoff (1969) to the moderately and severely mentally retarded This assessment paradigm consists of pre- and posttesting with an intervening session of task-relevant instruction on similar but nontest items. Budoff (1967) describes three types of response to the modified Kohs Block Design task among IQ-defined mildly retarded students: those who demonstrate a substantial increment as a function of the coaching experience (gainers), those who do not (nongainers), and those who score high on the pretest (high scorers).

In the first study, the Kohs block designs used by Budoff and Friedman (1964) were modified so that moderately and severely retarded students could successfully respond to the easiest items. Simple color discrimination and color matching items were added at the lowest level of difficulty. The next levels of difficulty included problems consisting of up to four one-color blocks. The most complex items presented patterns which required the students to manipulate four half-color blocks. These problems included designs similar to the easiest items of the standard Kohs series, and acted as a bridge to the designs used by Budoff and Friedman. Hamilton and Budoff (1974) reported that the trained students, as a group, did improve their competence on the task more than the non-trained group. Following training some students did markedly improve their scores, while others did not. Consistent with the findings for the mildly retarded, the moderately (IQs from 32 to 51) and severely (IQs from less than 10 to 31) mentally retarded gainers performed better on a simple learning task and were perceived as more effective learners by their classroom teachers than those who did not improve their score following training. The first purpose of this study was to obtain additional validity data for this test-train-test assessment approach with moderately and severely mentally retarded persons.

Hamilton and Budoff also observed that some subjects could correctly solve nontrained instances of problems during or immediately following training but failed to apply the strategy to equivalent test items during the posttest session on the following day. They seem to have acquired the solution strategy during training since they used it successfully with a minimal lapse in time, but they had difficulty retaining and/or spontaneously applying it after a one-day interval. Since the intent of learning potential assessment is to ascertain a person's ability to learn and apply problem-solving principles under optimal conditions, an alternative procedure was developed in which training is incorporated within the assessment process to allow the person to immediately apply the strategy to nontrained instances, rather than following the one-day interval upon retest. A second purpose of this study was to compare the responses of moderately and severely retarded persons with the training embedded within the testing procedure (train-within-test) or presented in a separate session (test-train-test).

Method

Subjects

The sample consisted of 38 mentally retarded persons drawn from the same institution as the previous study. From information drawn from the institutional records, the subjects ranged in chronological age (CA) from 12 to 22 years (mean CA = 204.5 months ±32) and in IQ from untestable up to the low 50s (mean mental age (MA) = 46.11 months ±15.9). As determined by responses to the easiest modified Kohs test items, the subjects' visual, motor, and color discrimination abilities were sufficient to perform the kinds of tasks used in this study. Length of institutionalization ranged from a minimum of one year to 20 years.

Instruments

Measurement of Learning Potential (Test-Train-Test Procedure). The test consisted of a modification of the Kohs block designs used by Budoff and Friedman (1964). Items were developed which constituted a logical downward extension of the Kohs designs, so that moderately and severely retarded students could successfully complete the easiest items. The test was composed of 13 difficulty levels with three test items at each level.

The test can be characterized as a match-to-sample task with block design pictures serving as the sample and concrete objects (one-inch [2.5 cm] blocks) available for construction of the match. Test problems ranged in difficulty from simple color discriminations to complex four block pattern constructions.

Thirteen demonstration-training items were developed to correspond to each difficulty level on the test. These items were used in the two testing sessions and the training session. During testing, these items were used by the examiner to demonstrate equivalent design constructions at particular difficulty levels. During the training session, the same items were used to teach the subject strategies for solving problems at a given difficulty level. Each subject was trained on four of these non-test items: the training item which corresponded to the highest level of difficulty (s)he successfully completed during pretraining, followed by the training items which corresponded to the next three difficulty levels.

The Extended Kohs Test was administered individually two times: prior to and one day after training. During each testing session the examiner constructed a demonstration item prior to each problem level. After viewing this demonstration, the subject was required to construct two of the three test items correctly from the designs to achieve credit at that problem level. Testing was discontinued when the subject failed to attain this criterion for three successive difficulty levels.

On the day prior to the second test administration an individual training session was interpolated which lasted no more than 30 minutes. The training procedure consisted of several steps: (a) assisting the subject initially in his construction of the design, (b) calling attention to the separateness of the blocks making up the design, (c) requiring the subject to actively point, block-by-block, to his/her constructed design and the corresponding "blocks" on the design card, and (d) repeating the process so that (s)he could practice and become familiar with the materials.

Training-within-test procedures were developed focused around the Arthur adaptation of the Leiter International Performance Scale and Raven's Progressive Matrices (Sets A, AB, B). Both tasks appeared suitable because they include problems which range over a broad span of difficulty, and can be responded to nonverbally.

The Arthur Adaptation of the Leiter International Performance Scale (Arthur, 1952) presents reasoning problems in a format that requires the subject to complete the problem by placing serially the appropriate one inch cubes, which contain problem-relevant stimuli printed on one surface, within spaces arranged in a specially designed tray. The test items range in difficulty from the two year level to adulthood. The response process requires no verbalization by the subject or the examiner after the initial orientation. For example, the tasks can be presented in mime to deaf persons. The Leiter taps a variety of skill areas including color matching, block design, shape and color discrimination, and seriation, to name a few.

The test procedure was modified so as to provide a correction and training procedure when an item was missed. The examiner immediately removed an incorrectly placed block, pointed to the correct one and encouraged the subject to correct his error. After this correction, the subject was given a maximum of three trials to complete the item with no assistance from the examiner. If the subject missed the item on any of the following trials, the correction procedure was repeated.

A weighted scoring system was used in which one point was awarded for each item correct on the first trial, .75 for each item correct on the second trial (corrected once), .50 for each item correct on the third trial (corrected twice), .25 for each item correct on the fourth trial (corrected three times), and .00 for each item that required more than four trials. Testing was discontinued after six items in succession were missed on the first trial.

The Coloured Progressive Matrices (Raven, 1956) were designed for use with children from the age of 5 through 12 years. Raven (1965) has described the test as a measure of "observation and clear thinking." He stated that the test "can be used satisfactorily with people who, for any reason, cannot understand or speak the English language, with people suffering from physical disabilities, aphasias, cerebral palsy, or deafness, as well as with people who are intellectually sub-normal or have deteriorated." (p. 3)

The Coloured Matrices (Sets A, AB, B) (1956) used in this procedure consist of three sets, each containing 12 items. Each problem consists of a design with one part missing. The examinee is required to choose one of six choices which best fits the missing part in the design. The test was modified in three ways. First, the items were presented in a concrete, manipulable form by mounting the patterns and choices on pieces of cardboard. Second, the test items were arranged in order of difficulty within the factor structures found by Corman and Budoff(Note 1): Simple Continuous Pattern Completion, Discrete Pattern Completion, Continuity and Reconstruction of Simple and Complex Structures, and Reasoning by Analogy. The order of presentation was: A1, A2, A3, A4, A5, A6, AB3, AB2, AB1, B1, B2, A8, AB7, B3, A7, AB4, B4, A9, A10,AB5, AB10, AB6, B5, AB9, AB11, B6, AB8, B9, B10, A11, B7, B11, B8, AB12, A12, B12. The purpose of presenting the items in this order was to maximize the probability that the training effect would transfer to subsequent, similar test items. Third, the test was modified by providing a correction and training procedure during the test for each item missed. The examiner removed the incorrect choice, placed the correct choice in the space, and pointed out the relevant features of the pattern in an attempt to demonstrate the appropriateness of the choice in completing the design. The correct choice was then removed and placed with the distractors, all of which were then rearranged randomly. The subject had three additional trials in which to complete the item correctly. If, however, the subject missed the item on the following trial(s) the correction procedure was repeated. The same weighted scoring system used with the Leiter was used in scoring this test. Testing was discontiuued after five items in succession were missed on the first trial.

The validity measures included two IQ tests (Stanford-Binet, Peabody Picture Vocabulary Test), a simple measure of learning, and ratings of student competence in classroom learning, and in dormitory behavior. The Stanford-Binet Intelligence Scale (Terman & Merrill, 1960) and the Peabody Picture Vocabulary Test (PPVT) (Dunn, 1959) were administered individually using the standardized instructions.

The Knox Cube Test (Arthur, 1947) is a nonverbal test which requires the subject to reproduce visually presented tapping sequences. Specifically, the examiner demonstrates a pencil-tapping sequence on four stationary wooden blocks, e.g., positions 1, 4, 3. The Examinee must attend, remember the sequence, and tap the blocks in the same order. The original scale includes items ranging from 2 to 7 elements within a sequence. It was modified by adding 6 items to the scale: 4 one-sequence taps to determine whether the subject had the idea of the task, and 2 two-sequence items to extend the lower end of the scale.

Each subject's classroom teacher was asked to indicate on the basis of his/her experiences with the student whether he/she tends to retain information or skills mastered in the classroom on a previous day.

The Resident's Functioning Questionnaire (Decker, Note 2) was completed by attendants familiar with the subject's abilities as observed in the ward environment. The scale consists of 22 items covering several areas: there are 7 items relating to self-help skills, 4 items on physical skills, 4 items on social skills, and 1 item for each of the following areas: language use, attention span, task performance, safety, destructive/assaultive beahvior, self-injury, and anxiety. The format for each item permitted the attendant to choose one of five descriptive statements.

The six instruments were administered by two trained examiners over a period of approximately two months to all subjects. All tests were administered to each subject in the same order, as follows: PPVT, Extended Kohs Test, Knox, Stanford-Binet, Leiter, and Raven's Matrices. Concurrently, the teachers' ratings and the Decker scale ratings were obtained from the subjects' teachers and ward attendants, respectively.

Results

The analyses relating to the validity of the Extended Kohs Test included the scores derived from the Raven and Leiter learning potential procedures so that comparisons would be available among all the variables.

Table 1, which presents the intercorrelations of all variables, indicates that following K1 (Extended Kohs pretest), the zero-order predictors of K2 (Extended Kohs posttraining scores) in order of magnitude were the weighted Leiter score (.82), Knox Cubes (.71), Binet mental age (.61), weighted Raven score (.59), teacher rating (.51), PPVT mental age (.49), dormitory ratings of behavioral competence (.47), and Stanford-Binet IQ (.39). All were significantly correlated with K2 above the .05 confidence level. The moderately and severely mentally retarded person's response to the training experience on the block design task clearly provides an indication of the level of competence the person displays in learning, whether measured by laboratory tasks, classroom performance, or estimates derived from dormitory behavior. These results provide support for this measure of learning potential with these kinds of persons.

Insert Table 1 about here

The intercorrelation matrix was then used in a multiple regression analysis, employing the least squares solution, with the posttraining Kohs

Table 1

Intercorrelations of Experimental Variables (N = 38)

	1	2	3	4	5	6	7	8	9	10
1. SB MA	---	.787***	.830***	.702***	.257	.518***	.717***	.565***	.485**	.609***
2. SB IQ		---	.558***	.467**	.191	.160	.473**	.380*	.428**	.393*
3. PPVT MA			---	.569***	.050	.442**	.600***	.426**	.297	.491**
4. K1 (pretraining)				---	.454**	.620***	.820***	.635***	.511**	.904***
5. Teacher rating					---	.261	.466**	.606***	.350*	.505**
6. Raven (weighted)						---	.712***	.411*	.426**	.590***
7. Leiter (weighted)							---	.519***	.549***	.824***
8. Knox								---	.410*	.710***
9. Decker									---	.469**
10. K2 (posttraining)										---

*p <.05

**p <.01

***p <.001

score (K2) as the dependent variable. Independent variables in the equation on K2 were the remaining array of measures. The order of entry of independent variables in the stepwise regression was by the amount of variance accounted for by each variable.

The regression analysis on posttraining scores indicated that the significant predictors of K2, with the variance due to K1 removed, were the Knox and the Leiter scores. Although in this equation the Raven was not a significant predictor of a subject's responses to training, inspection of the zero-order correlations reveals that the Raven's score was highly correlated with K1, Knox, and Leiter, and these four independent measures were significantly related to K2. Since a large portion of the variance in the posttraining score (K2) was shared among K1, Knox, and Leiter, the Raven's Matrices were not a significant predictor of K2 with the other performance tests entered into the equation. As indicated by the rationale for a training-based learning potential measure and consistent with previous findings (Budoff & Corman, 1974), the simple learning task and the learning potential scores, both of which provide practice and positive feedback (unlike the measures derived from the traditional intelligence test procedure) constituted the significant predictors of the Kohs posttraining score, when the variance attributable to the pretrained competence was removed.

Insert Table 2 about here

To summarize performance on the Raven and Leiter train-within-test learning potential procedures, means and standard deviations were calculated across all subjects on total items attempted, items correct without

Table 2

Raw Score Means and Standard Deviations of Total, High LP, and Low LP Groups on the Raven and Leiter

		Items attempted		Items correct without training Trial 1		Additional items correct						Total correct (four trials)	
						Trial 2		Trial 3		Trial 4			
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Raven													
Total	38	23.13	8.70	11.71	6.71	7.39	3.77	2.07	1.36	.79	1.08	21.97	8.83
High LP	21	24.57	8.34	13.90	6.94	7.57	3.69	2.10	1.41	.48	.59	24.05	8.27
Low LP	17	21.35	8.80	9.00	5.28	7.18	3.87	2.06	1.30	1.18	1.38	19.40	8.84
Leiter													
Total	38	22.32	6.39	11.74	6.06	3.39	1.84	1.87	1.40	.79	.92	17.79	7.11
High LP	21	25.81	5.74	15.05	4.86	3.86	1.70	1.95	1.43	.90	1.02	21.76	5.28
Low LP	17	18.00	4.10	7.65	4.75	2.82	1.85	1.76	1.35	.65	.76	12.89	5.92

training (Trial 1), items correct with successive training (Trials 2, 3, 4), and total items correct over four trials. To compare these persons' performance on these train-within-test procedures, the subjects were dichotomized on the basis of their performance on the Kohs test-train-test procedure into high and low learning potential groups. Table 2 presents those data as well as means and standard deviations for high and low learning potential (LP) groups. High LP was defined as an increment of two or more difficulty levels on the Extended Kohs Test following training or attainment of difficulty level 4.0 or higher on Extended Kohs pre- or posttest.

The results indicate that as a total group the subjects attained approximately half the items attempted with no training (Trial 1) and substantially benefited from training. They responded correctly, over all four trials, to 93% and 78% of the Raven and Leiter items, respectively. It is interesting that on both tests most of the increment occurred after only one training experience (Trial 2), and that the pattern was consistent for high and low learning potential groups. Converting the mean Raven raw score for the total group before training (11.71) and after successive training (21.97) to corresponding mental ages yields 5.5 and 8.16 years, respectively. Leiter raw scores before (11.74) and after training (17.79) convert to mental ages of 4.5 years and 5.0 years respectively. For both tests, the mental age equivalents are higher than the PPVT and Stanford-Binet mental age estimates of approximately 4.0 years. High Kohs LP persons were correct on a greater number of items prior to training on the Raven's problems than the low Kohs LP group ($t = 2.47$, $p < .05$). None of the other differences between high and low

LP groups on the Raven were significant. On the Leiter, the high LP group performed significantly higher than the low LP group on number of items attempted ($\underline{t}$ = 4.88, $\underline{p}$ <.001), number of items correct without training ($\underline{t}$ = 4.73, $\underline{p}$ <.001), and total items correct over four trials ($\underline{t}$ = 4.82, $\underline{p}$ <.001).

Discussion

The results of this study provide further support for the findings of Hamilton and Budoff (1974) that a learning potential approach to the assessment of abilities of moderate and severely retarded persons provides a fairer estimate of their capabilities than traditional IQ measures. Since these persons lead markedly atypical lives and experience the world quite differently from the persons on whom the standardized tests were normed, valid attempts to measure their abilities must, at a minimum, insure that these persons have (a) the physical capabilities to perform the task and (b) some direct or related experience(s) with the task in question. Due to the interaction of the distortions in their life experience and the biological complications that accompany moderate and severe levels of mental retardation, these persons frequently do not sustain their attention for extended periods, nor do they critically attend to the distinctive features of a stimulus display or spontaneously develop a work-oriented stance to problems set before them.

The introduction of training in the assessment procedure changes a number of features of the test situation for the participant. The most evident change is affective. The person changes his perception of the situation from a purely evaluative one, with the negative affect he attaches to such situations, to one in which he perceives the role of the examiner as helping him become competent. This change was reflected in the evident pleasure with which these persons approached the learning potential assessment sessions. While the social reinforcement value of the contacts (Zigler, 1966) must be considered, these persons enjoyed the contacts and came repeatedly without qualms, though their initial response to tests usually evokes negative affect and avoidant behaviors. The active intervention of the examiner in his role as trainer helps the person learn to focus on the critical features of the stimulus display, helps him to structure an attack on the problem, and encourages the person to actively test his hypotheses regarding the critical features of the problem without fear of the negative judgment. The procedures stimulate and sustain an active work-oriented address to these problems, which probably can be extended to other situations when this type of context for learning and doing is established.

These findings with learning potential assessment provide support for Schucman's (1960) statement that learning scores with these types of persons are stable and consistent indicators of ability to learn and profit from experience. The findings also are in essential agreement with Gold (1972) who has reported similar results with such manipulative tasks as bicycle brake assemblies. Training on a simple brake assembly resulted in more rapid and successful learning of a more complex brake assembly and in more rapid mastery of an electronic circuit board assembly.

Within the context of this study, none of the commonly used measures (Stanford-Binet mental age, Stanford-Binet IQ, PPVT mental age) were positively correlated with the teachers' ratings of these persons' learning and retention behavior. By contrast, posttraining scores were correlated with the teachers' ratings, and the dormitory ratings of behavioral competence in their living situations. In two "real life" situations, persons who have extended contacts with these mentally retarded individuals perceive ability differences in coping with their environment that are also described by posttraining scores. These findings provide further support for a learning potential assessment in which the judgment of ability to perform is made following training, rather than following a single test administration. This finding suggests that teacher and attendant judgments about these persons' abilities are based on skills different from those measured by the PPVT and Stanford-Binet.

This study failed to provide empirical evidence to support a choice of either the test-train-test or train-within-test assessment formats. The data provided support for a training-based model, regardless of the format utilized. Both formats were equally effective when evaluated against the predictive criteria.

Intuitively, a training-embedded-within-test format should overcome some of the problems that appeared to be evident on the test-train-test (on separate days) paradigm. In the latter procedure, test-related anxiety does appear during the pretraining session, which, however, gets handled within the train-within-test procedure at the first point at which the child errs. More important, the individual can demonstrate his ability to master and apply a principle within the context of the ongoing assessment unlike the test-train-test format when the principles he learns must be retained to be applied at a subsequent session.

In the train-within-test format, transfer of the newly acquired understanding is evident on successive items within a series utilizing similar principles during the same session, maximizing the opportunity to see the person's ability to utilize the newly acquired principle of solution with nontrained items. In the present context of concern with providing maximal opportunities for these individuals to progress, this evidence of ability to learn and apply that understanding provides a sensitive index of his trainability that should influence the type of program developed for the individual. If one wishes, one could re-administer the train-within-test procedure at some subsequent date to ascertain the staying power of the newly acquired understandings and so test for longer term transfer.

Both learning potential procedures permit one to view the individual's response to training. However, the train-within-test format allows the tester to quantify (across trials) the amount of assistance the child requires to reach a criterion performance on given items. Different patterns of response to the training become evident. Some of these persons require considerable help on the initial problems, i.e., needing three to four trials on the problems presented early in a series. After some such exposures, they solve more difficult instances of these problems on the initial trial or after one set of prompts.

Other persons may not profit from the instruction at all, as reflected in their continuing need for three or four trials after which they may or may not demonstrate their understanding of the problem. The training procedure can also indicate which type(s) of correction procedures the different persons may respond to: verbal prompts, pointing, modeling, or kinesthetic tracing. If these types of prompts can be considered in hierarchical order, does the individual who requires the kinesthetic prompts, initially, respond at a later stage in the training to the higher order prompts? These patterns of response to the instruction in this micro-teaching situation can provide meaningful information for the teachers and trainers of these individuals when their education and training plan is being formulated.

Further research is required to understand more clearly the manner in which the information acquired by this learning potential assessment procedure can be translated into suitable plans for education and training of these persons.

Reference Notes

1. Corman, L., & Budoff, M. Factorial composition of the Raven Progressive Matrices in the context of learning potential assessment. RIEP-print No. 62, 1973, Research Institute for Educational Problems, 29 Ware Street, Cambridge MA 02138.
2. Decker, H. A system for planning and achieving comprehensive health care in residential institutions for the mentally retarded. Ann Arbor: Department of Health Development, School of Public Health, University of Michigan, 1970.

Footnote

This research was supported by grant OEG-0-8-080506-4597 from the Bureau of Education for the Handicapped, U.S. Office of Education, and grant NE-G-00-3-0016 from the National Institute of Education, U.S. Department of Health, Education, and Welfare. James L. Hamilton is now at the Bureau of Education for the Handicapped, U.S. Office of Education. The Research Institute for Educational Problems is located at 29 Ware Street, Cambridge MA 02138.

References

Arthur, G. A point scale of performance tests. New York: Psychological Corporation, 1947.

Arthur, G. The Arthur adaptation of the Leiter International Performance Scale. Washington: Psychological Service Center Press, 1952.

Barrett, B. H. Acquisition of operant differentiation and discrimination in institutionalized retarded children. American Journal of Orthopsychiatry, 1965, 35, 862-885.

Budoff, M. Learning potential among institutionalized young adult retardates. American Journal of Mental Deficiency, 1967, 72, 404-411.

Budoff, M. Learning Potential: A supplementary procedure for assessing the ability to reason. Seminars in Psychiatry, 1969, 1, 278-290.

Budoff, M., & Corman, L. Demographic and psychometric factors related to improved performance on the Kohs Learning Potential procedure. American Journal of Mental Deficiency, 1974, 78, 578-585.

Budoff, M., & Friedman, M. "Learning Potential" as an assessment approach to the adolescent mentally retarded. Journal of Consulting Psychology, 1964, 28, 434-439.

Clarke, A. M., & Clarke, A. D. B. Learning transfer and cognitive development. In J. Zubin & G. A. Jervis (Eds.), Psychopathology of mental development. New York: Grune & Stratton, 1967.

Dunn, L. M. Peabody Picture Vocabulary Test manual. Minneapolis: American Guidance Service, 1959.

Gardner, L. P. Response of idiots and imbeciles in a conditioning experiment. American Journal of Mental Deficiency, 1945, 50, 59-80.

Gold, M. W. Stimulus factors in skill training of retarded adolescents on a complex assembly task: Acquisition, transfer, and retention. *American Journal of Mental Deficiency*, 1972, *86*, 517-526.

Hamilton, J. L., & Budoff, M. Learning potential among the moderately and severely retarded. *Mental Retardation*, 1974, *12*, 33-36.

Jensen, A. Learning ability in retarded, average, and gifted children. *Merrill Palmer Quarterly*, 1963, *9*, 123-140.

Luria, A. R. An objective approach to the study of the abnormal child. *American Journal of Orthopsychiatry*, 1961, *31*, 1-14.

O'Connor, G., Justice, R. S., & Payne, D. Statistical expectations of physical handicaps of institutionalized retardates. *American Journal of Mental Deficiency*, 1970, *74*, 541-547.

Penrose, L. S. *Mental defect*. New York: Farrar & Rinehart, 1934.

Schucman, H. Evaluating the educability of the severely mentally retarded child. *Psychological Monographs*, 1960, *74*, 1-34 (Whole No. 501).

Silverstein, A. B. Psychological testing practices in state institutions for the mentally retarded. *American Journal of Mental Deficiency*, 1963, *68*, 440-445.

Talkington, L. W. Over-achievement or stimulation? A case study. *Digest of the Mentally Retarded*, 1967, *4*, 21-23.

Terman, L. M., & Merrill, M. A. *Stanford Binet Intelligence Scale*. Boston: Houghton Mifflin, 1960.

Tobias, J. The evaluation of the vocational potential of mentally retarded young adults. *Training School Bulletin*, 1960, *56*, 122-135.

Vygotsky, L. S. *Thought and language*. Cambridge, Mass.: M.I.T. Press, 1962. (Originally published, 1934.)

Zeaman, D., & House, B. J. The role of attention in retardate discriminition learning. In N. R. Ellis (Ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963.

Zigler, E. Mental retardation: Current issues and approaches. In L. W. Hoffman & M. L. Hoffman (Eds.), Review of child development research (Vol. 2). New York: Russell Sage Foundation, 1966.

PRINTS

FINAL REPORT:
AN EDUCATIONAL TEST OF THE VALIDITY OF LEARNING POTENTIAL MEASUREMENT

Elisha Babad and Joseph Bashi
School of Education
Hebrew University of Jerusalem

RIEP-Print Number 91
June, 1975

RIEP-Prints are a continuation of the series STUDIES IN LEARNING POTENTIAL

RESEARCH INSTITUTE FOR EDUCATIONAL PROBLEMS, INC.
29 WARE STREET • CAMBRIDGE, MASSACHUSETTS 02138 • TELEPHONE (617) 868-0360-2

© 1975

ABSTRACT.

In the Learning Potential (LP) tests, conceptual (Level II) ability is measured in a "training-within-test" paradigm. In this study we compare the validity of LP, verbal IQ and standardized achievement in predicting attainment is a specially designed curriculum in cryptarithmatic. The program required minimal prerequisites of the students, progress in the program was graded in a multiple support system, the role of memorization and verbal expressive usage was minimized, all in an attempt to bring all students to a level of competence in solving level II type cryptarithmetic problems. Two hundred students in eight seventh-grade classes studied the program for 15-20 class sessions and were then tested.

The results show that the majority of the students attained a level of competence in this program, with no differences between groups differing in educational home background. The LP posttest score was found to be superior to all variables in predicting success in that program. It's superiority to the other predictors was especially marked for the disadvantaged segment of the sample.

Chapter 1: Introduction - The LP conceptualization.

Children from poor and/or nonwhite homes tend to score at below-average levels on tests which purport to measure intelligence. Jensen (1969) has argued that the mean difference of 15 IQ points between white and black groups represents a real difference in inborn general ability. However, the IQ difference has also been frequently explained by handicaps that poor children bring to the testing situation (Hunt, 1961). They are fearful of the testing process, expect to do poorly, are often insensitive to speed requirements, are unfamiliar with the problem contents, and do not develop spontaneously the most effective strategies (by middle-class criteria) to solve the problems.

IQ tests measure the degree to which children have spontaneously acquired from their natural environment, the skills and knowledge which cumulatively predict academic school success. The plausible assumption is made that a child who was learning informally prior to entering school will continue to learn - formally and informally = in and out of school. Children from non-middle-class homes who do not have an equal and frequent access to school-preparatory experiences tend to score poorly on IQ tests and are often viewed as "less intelligent".

Yet, many of these low-IQ children are competent problem-solvers in their non-school environment, having mastered the skills, knowledge, and strategies necessary to maintain a successful adjustment. These children do learn and profit from relevant experiences more than their IQ scores and school achievements indicate. In Hunt's (1961) terms, this difference in competence may represent the problem of the match between the environmental demands of the school and its tests, and the child's existing schemata in his familiar world. Assessment procedures must be developed which will optimize the match and result in more culture-fair measurement of general ability.

Eels, Davis, Havighurst, Herrick, and Tyler (1951) developed problem games that were deemed relevant to the experiences of low-socioeconomic-status children, but found no change in the gap between the different social class groups. This may have been due to the demand that the child go about solving the problems as middle-class psychologists thought he should. Changing the contents of the problems was not sufficient to enable the lower-class children to narrow the gap between their problem solving styles and those expected by the middle-class bias of the test constructors.

Budoff (1969) has described a learning potential (LP) procedure for assessing ability among low-IQ, low-socioeconomic-status children based on a process-oriented conceptualization of intelligence. The focus of the LP paradigm is on the child's trainability, that is, his ability to improve performance on reasoning problems following a systematic learning experience. Reasoning is viewed as the critical ability and the LP paradigm focuses on the ability to learn new strategies and apply them to relevant reasoning problems. The reasoning tasks are administered in a "test-coach-test" sequence. Training allows the low-SES child to understand how to solve the problems when the contents of the problems may initially be strange and the appropriate strategies not readily apparent to him.

Training is hypothesized to be particularly critical for the child from a poor background, who may learn different cognitive strategies in different expressive formats than those presumed to be available by the tests The training helps the child to narrow the cognitive gap between his previously learned problem-solving strategies and those implicit to the problems he must ordinarily solve on the tests in the middle-class school. The repeated contacts with the materials in a context of support and success allow the school-failing but able child to assert his sence of competence. Without this competence boost, he tends not to perform at his best, implicitly expecting failure (Zigler, 1966). The essence of this assessment method then, is to impose some control on the potentially negative effects of prior life experiences on his test performance.

Naturally, no en bloc improvement in performance is expected, but rather a differential change reflecting the children's potential. The LP paradigm is thus assumed to tap both problem-solving ability (actual and potential) and the ability to learn new problem-solving strategies and to apply them efficiently to relevant problems.

In the LP tests, the pretest score reflects the subject's present level of functioning and his existing ability to work with the problems (in that sense, the LP pretest can be part of any IQ test). The training is designed to bring all subjects to an equalized starting point, providing them with an opportunity (in a context of success and support) to learn and practice the relevant strategies for handling the problems at hand. Posttraining scores thus reflect the child's potential ability to learn and perform under optimized conditions.

The problems employed in the learning potential tests fall within Jensen's (1969) level II category (conceptual learning). According to Jensen, there is little overlap between the level II curves of lower-and higher-SES groups. Jensen's theory is based on the spontaneous responses evoked in traditional "one shot" intelligence testing. According to the LP conceptualization these testing procedures underestimate the potential level II ability of great number of disadvantaged children, while many middle-class children develop efficient strategies even without direct training, thus showing less change from pre- to post-test performance. One could therefore expect to find greater overlap between the curves of lower and higher-SES groups on LP posttests than on LP pretests and other traditional IQ tests. Complete overlap between the curves is not assumed. Inclusion of training procedures within ability testing may thus identify the false positives among the disadvantaged (Ortar, 1960). This hypothesis was confirmed by Babad and Bashi (1975) in a recent study. Posttest curves of disadvantaged groups show greater overlap with the curve of middle-class children than the pretest curves.

The date, only nonverbal reasoning tasks have been used in measurement of LP, since these tasks are not dependent on language and verbal expressive skills, which represent another major area of difficulty for disadvantaged children. Budoff (1969) used an altered version of the Kohs Block Designs; Jacobs (1971), Linn (1970), Feuerstein (1968) and Budoff, Corman, and Litzinger (1973) used different versions based on Raven's progressive matrices; Ortar and Babad (1973) and Babad and Budoff (1971) used series completion tasks. In all cases, coaching took place between repeated administrations of the task, providing the students with the opportunity to learn appropriate problem-solving strategies and an occasion to understand the demands of the task at hand in a relaxed atmosphere. Efforts were made in Budoff's laboratory to construct and validate Learning Potential Tests for group administration at all age levels. In the present study, we used revised versions of the Series Learning Potential Test (SLPT) and the Matrices Learning Potential Test (MLPT).

Validity of LP:

Ability tests constructed for use by educational systems have usually been validated in terms of their power to predict school achievement.

Conceptually, both achievement tests and the LP tests are based on the premise that testing should follow relevant tuition, and that the student cannot be expected to perform at his highest level until he has been provided with all necessary preparatory experiences. The conventional intelligence test is implicitly based on the same assumption, only that it deals with indirect learning such as the acquisition of vocabulary and general information. Indeed, many disadvantaged children do not have such opportunities, and therefore their potential level of ability is underestimated by their IQ scores.

Despite the conceptual similarity between LP and achievement, the former cannot be expected to be superior predictor of school achievements due to the following reasons: achievement tests lean heavily on the verbal expressive mode while LP is non-verbal; they test memory and skills as well as reasoning; and the tests are administered to children who have not been brought to an equalized starting point and therefore differ significantly from the very start.

In the last 30 years, IQ tests have undergone changes designed to increase their "school relevance" (i.e. to improve their validity in predicting standardized school achievement). In this way, the IQ test has became a more effective empirical predictor of school achievement. However, this validity is partially based on bias which is common to both tests. In other words the very same factors (e.g. language skills, cumulative deficit) which depress the performance of disadvantaged children on IQ tests also affect their regular school performance, yielding accurate, but biased predictions.

Budoff, Gimon and Corman (1974) tested 188 Spanish - American children with the SLPT, the MLPT, the WISC and the Inter American General Abilities Series (IAGAS). In a multiple regression analysis using the achievement scores as criteria, both the SLPT posttest score and the WISC Performance score were found significant predictors (approximately of the same potency) of the nonverbal and numerical achievement scores. The WISC Vocabulary score (the only verbal test available in Spanish) and the MLPT posttest score significantly predicted one of the verbal achievement scores.

Corman and Budoff (1973) tested almost 1,000 students in New England with the SLPT and the Gates-MacGinitie achievement test. The correlations between the SLPT posttest score and the Gates-MacGinitie scores ranged from .43 to .59, a range typical and acceptible for most IQ tests. Unfortunatly, no IQ test was administered in this study.

Sewell (1973) tested 62 first grade black children of low-SES with the MLPT, the WISC, and the Stanford Achievement Test Primary I Battery (Kelley, 1964). While predictions of reading and arithmetic scores from the MLPT pretest score were rather weak (.22 and .25), the MLPT posttest scores were much stronger predictors (.47 and .54, respectively), respectively), reaching the validity level of the WISC Full Scale IQ score, and surpassing the separate predictions of the Verbal and Performance IQ scores.

Babad and Budoff (1971) compared the validity of the SLPT and the TOGA IQ (Flanagan, 1960) in predicting Wide Range Achievement Test scores for 60 EMR (educable mentally retarted) grade school children placed in special classes. The SLPT posttest was superior to the TOGA IQ in predicting arithmetic scores, equal to the IQ in predicting reading scores, and inferior to the IQ in predicting spelling scores. The SLPT posttest was superior to IQ also in predicting teacher ratings of general ability and academic potential (.52 and .50 for the SLPT, compared to .46 and .44 for IQ).

Babad and Budoff (1974) compared the validity of SLPT and TOGA IQ scores in predicting teacher ratings of 207 third grade children in a New England town. The teachers rated the children for achievements in reading, spelling, arithmetic, general achievement, and academic potential. The correlations reported by Babad and Budoff show no difference in predictive validity between the SLPT and the TOGA IQ variables.

It is evident from these findings that school achievement, measured either by standardized tests or by teacher evaluations, can be equally well predicted by standard IQ scores and by the LP posttest scores. Babad and Budoff (1974) argue that since the target population of the LP tests in the low-SES, disadvantaged group, the validity of these tests should be examined on disadvantaged samples as well as on samples representing the total range of social class in society. They hypothesized that training would increase the sensitivity (and subsequently, the validity) of the SLPT for low-IQ (low-achieving) students. To test this hypothesis, they divided their sample (n=207) into three subgroups: Subnormal (labeled in Massachusetts "Educable Mentally Retarted (EMR)", IQs below 80); Dull-to-average (IQs between 80 and 99); and Bright-normal (IQs of 100 and above). Correlations of SLPT and TOGA scores with the five measures of teacher rated achievement (reading, spelling, arithmetic, general achievement, academic potential) were calculated separately for each group. Correlations for the five ratings in each cell were averaged, and they are presented in Table 1. As could be expected in a case of a restricted range, all subgroup correlations were depressed. However, the LP predictions were superior to the IQ predictions,

Table 1. Predictions of teacher-rated achievement from TOGA IQ and SLPT scores for three ability groups.

		Subnormal ($IQ < 80$, $n = 36$)	Dull-to-average ($80 \leq IQ \leq 99$, $n = 95$)	Bright-normal ($IQ > 100$, $n = 76$)
IQ	Verbal	.11	.08	.27
	Reasoning	.20	.19	.25
	Total IQ	.16	.15	.31
SLPT	Pretest	.29	.30	.37
	Posttest	.35	.34	.35

In all groups, and particularly in the low-ability group, in the same report, Babad and Budoff showed that the two lower groups gained more than the high-IQ group from training in problem-relevant strategies, and the dull-to-average group gained more than the other groups from repeated administrations without training. In addition, they reported that substantial proportions of subnormal (EMR) subjects reached the average reasoning level of their non-retarded peers following the short training session.

Thus far, we have brought evidence bearing on the predictive validity of LP and IQ when the criterion, school achievement was accepted as conventionally measured (by standardized tests and teacher ratings).

As we have argued earlier, both IQ and standardized achievement are oriented toward measuring the accumulated acquisition (formal and informal) from past experiences and both may be biased towards the child who has been less exposed to the relevant experiences. The LP paradigm, diagnostic by design, intentionally neglects past acquisition by ensuring that all children are brought to an equalized starting point in terms of their preparatory experiences, thus predicting their potential level of future functioning under optimal conditions. School systems all over the world are becoming more and more flexible in fitting their demands to the individual student in order to maximize attainment and to narrow, as much as possible, the gap between demands and ability. This requires a more dynamic approach to the construction of curricula and achievement tests, so that they could enhance the chances of all potentially able children (middle-class and disadvantaged) to reach higher levels of performance. Such curricula would be process-oriented (rather than skill-or knowledge-oriented), would offer proper tuition to all children and would minimize the role played by accumulated deficit. In the present study, we examine the validity of LP and IQ in predicting achievement in a program designed according to these lines. Next we shall present previous projects related to the same issue.

The "Batteries and Bulbs" studies.

Budoff, Meskin, and Harrison (1971) designed a laboratory science program in electricity. The program makes minimal verbal demands, reduces the need for developed skills in reading and formal explanation. Instead, the program emphasizes learning by manipulation of materials. In the "Batteries and Bulbs" program the child experiments with his own equipment (flashlight batteries, bulbs, and various types of wires) and draws conclusions based on the results of his own observations in these experiments. The child learns, among other things, about different types of circuits, resistance, conductability and insulation. Few formal electrical terms are learned, and formal explanation of concepts is minimized . No attempt is made to offer a comprehensive theory of electricity. The final test in the program is divided in two parts: in the first part the problems are presented as actual electrical set-ups mounted on masonite pegboards, and the children empirically demonstrate their comprehension of the concepts and their application via direct manipulation. In the second part the problems are schematically presented in diagrams.

The subjects in this study were special class, EMR, junior- high-schoool students. According to their pretest and posttest performance on the KOHS LP assessment procedure they were divided into three groups: nongainers, gainers, and high scorers. Control groups were nontaught EMRS who did not participate in the program and regular class students who participated in the program. The results showed that LP accounted for differences in learning on the program more significantly than either IQ or regular vs. special class assignment. Gainers and high scorers showed higher performance than the nongainers.

Gimon and Budoff (1974) adapted the "Batteries and Bulbs" program to Spanish and administered it to Spanish-American children in several grades of the elementary school. The final electricity test was again divided into a verbal and a non-verbal test.

The results of the multiple regression analyses showed that the IQ variables did not significantly predict success in the program, while the LP posttest scores (especially the combined score of the MLPT and the SLPT) significantly predicted performance both in the verbal and the non-verbal tests. For the verbal electricity test, the significant predictors in the multiple regression analysis were age, the combined LP posttest score, and achievement tests in both Spanish and English. For the non-verbal electricity test, the significant predictors were age and the combined LP posttest scores, but not the achievement tests. The WISC scores were not found significant predictors in either analysis.

These data support the case of the validity of the LP measurement in non-middle class, culturally different populations. The gap between the validity of the standardized achievement scores in predicting verbal and non-verbal electricity attainments and the lack of validity of IQ in predicting both types of attainment are quite dramatic, supporting the claim that teaching programs and tests constructed to measure children's ability and achievement should be changed in order to identify the potentially able, disadvantaged children and facilitate the realization of their potential in appropriate compensatory curricula.

Chapter 2: The "Letters and Numbers" program

Following the basic design of the Batteries and Bulbs projects, we investigated the validity of LP for an Israeli population utilizing a program that differs from the Batteries and Bulbs program in the following aspects:

1) Prerequisite knowledge for entering the program should be minimal and mastered by all children.

2) The program should be more central to the core areas of the elementary school curriculum, both in its content and required processes.

3) The program should be mostly process-oriented, emphasizing processes such as hypothesis testing, application of problem-solving strategies and level II thinking.

4) The program should be verbally-based, but the verbal requirements should be minimal.

5) The program should be taught in the regular context of the classroom without disruption of the arrangement of the class, so that the program could be implemented in the entire school system.

6) The program should be inexpensive, with no need to provide each student with a kit containing materials and tools,and not necessarily focus on manipulation of materials.

The type of task selected for the project was "cryptarithmetic" (Newell and Simon, 1972). In a cryptarithmetic problem each digit in an arithmetical exercise is represented by a letter and the student must identify all digits and re-translate the exercise to numbers, using strategies and testing hypotheses based on several simple arithmetical rules.

We decided to limit the program to addition problems only, so as to minimize the number of rules that must be mastered for solution. The classic addition problem is:

$$\begin{array}{rl} & \text{G E R A L D} \\ + & \text{D O N A L D} \\ \hline & \text{R O B E R T} \end{array} \quad \text{(given the D=5)}$$

In order to solve this problem, one must be able to solve simple addition problems and know several simple arithmetical rules (e.g. that the carryover can not exceed 1, and that a number plus itself yields an even sum). . On the basis of these elementary prerequisites one must function on a high conceptual level, systematically searching for entry cues, using all relevant rules, applying strategies and testing hypotheses. Such problems and the processes required for their solution are relevant objectives to the school curriculum, and yet most teaching falls short of developing such processes.

In designing the program, we tried to achieve the following goals:

1) To minimize the role of frontal teaching, and maximize differential support by the teacher.

2) To actively involve the student with maximum stimulation

3) To provide attractive tasks and stimuli that would maintain the student's interest and motivation in any given session.

4) To make the program very gradual, so that all students could reach the criterion of each segment in a reasonable time.

5) To control the number of bits of information introduced in each segment and to avoid loading the students with information that is not strictly necessary for completing a a given task.

6) To keep the program on a minimally verbal level.

7) To avoid the creation of time pressures on the slower students by allowing individual pacing in a multiple support system.

8) To teach the program in the regular context of the classroom, but replace frontal teaching with a carefully programmed self study method.

Description of the program

The program consisted of 25 work sheets. Sheets were separated, so that each child had in his possession only one sheet at a time. Teacher explanations sometimes preceeded the introduction of work sheets, but the teachers' main function was to supervise the work and provide support when necessary.

The program was divided into three parts:

1) Learning and practicing the technique of cryptarithmetic-- 4 work sheets, 1-2 hours.

2) Learning and practicing the various rules necessary for solving cryptharithmetic addition problems. This is the major part of the program, consisting of 12 work sheets, requiring 8-10 hours.

3) Practice exercises, without new learning --9 work sheets, 3-5 hours.

In the first part the students practice transformations from numbers to letters and from letters to numbers. They must first learn that in a given problem every digit is represented by one letter which was arbitrarily determined. They also learn that the letter-digit correspondence differs in each problem. At this stage students are given two of three elements (1) the exercise in digits; 2) the exercise represented in letters and 3) the transformation key) and they must fill in the missing element. Little actual calculation is demanded at this stage. Exercises are designed to be attractive, and most transformations make up a funny combination of words when solved. Funny word combinations were used throughout the program.

In the second part of the program the various rules are learned and practiced. In the early stages, each rule is practiced separately. In the more advanced stages a given problem may require the use of several rules at once. Each rule is also explicated to give it "cue property" as entry point into solving a new exercise. The following rules are learned:

1) The sum of A+A is always even.

2) The sum of A+A or A+B never exceeds 18. Therefore, the carry-over can only be 1, and the sum with a carry-over will never reach 20.

One example of the cue property derived from this rule is the following:

$$\begin{array}{r} A\ M \\ +\quad I \\ \hline G\ O\ D \end{array}$$

In this exercise G=1.

3) When

$$\begin{array}{r} .\;.\;.\;A \\ +\;.\;.\;.\;A \\ \hline .\;.\;.\;B \end{array}$$

and B.is known (e.g. 4), thein A could be one of two digits (2 or 7) separated by a difference of 5, and depending on whether there is a carryover.

4) When

$$\begin{array}{r} .\;A\;. \\ +\;.\;A\;. \\ \hline .\;A\;. \end{array}$$

then: a) A=0, or b) A=9, if there has been a carryover.

5) When

$$\begin{array}{r} .\;A\;. \\ +\;.\;B\;. \\ \hline .\;A\;. \end{array}$$

then a) B=0, or b) B=9, if there has been a carryover.

Rules 4 and 5 have some features in common: the possible digits to be identified are either 9 or 0, and both rules are derived from the fact that X + 0 and X + 10 yield the same digit X. In both cases if the A + A = A or A + B = A would appear as last digits in the exercise:

$$\begin{array}{r} .\;.\;A \\ +\;.\;.\;A \\ \hline .\;.\;A \end{array} \qquad \text{or} \qquad \begin{array}{r} .\;.\;A \\ +\;.\;.\;B \\ \hline .\;.\;A \end{array}$$

Then the letter to be identified (A and B, respectively) can be only 0. The hypothesis that the letter could represent 9 should be tested only when these combinations appear in the middle of an exercise and there is a reason to expect a carryover.

6) In the exercise

$$\begin{array}{r} A\;M \\ I \\ \hline G\;U\;Y \end{array}$$

A = 9 and U = 0

In this part of the program, each rule is practiced separately with a variety of exercises. Later, two or more rules are needed to solve a given problem. The problems range in the amount of unknown information. The search for cues which constitute "entry points" is stressed at all times and the students are encouraged to note down every cue and every letter that was identified. In the early parts, explicit instructions are given as to how the steps should be sequenced. Later, these instructions are dropped. Students are trained to numerically check their solution before they move to the next problem.

The third part of the program consists of practice exercises with no explanation. However, the teachers can add any amount of support and explanation according to the situation in the class. Frontal presentations on the board are also permitted.

The exercises range in difficulty and in the number of rules required for solution of each problem. The following exercises represent a simple one-rule problem and a complex multi-rule problem, respectively:

$$\begin{array}{r} M\ M \\ \underline{\quad M} \\ B\ O\ S \end{array} \qquad \begin{array}{r} V\ O\ W \\ \underline{W\ I\ D\ O\ W} \\ D\ R\ O\ I\ R \end{array}$$

General objectives of the project:

1) The student population in the Israeli school system is very heterogeneous, due to the large immigration of groups varying widely in their cultural backgrounds. Over 30% of Israeli school children are officially classified as "disadvantaged". It is therefore crucial to investigate the efficacy of any instrument (test or curriculum) which might enhance the realization of potential of low-background-children.

2) To design construct, and test the efficacy of a learning program in cryptarithmetic which follows the goals presented above (i.e. reaching a high level of thinking based on minimal knowledge in a content area relevant to the school curriculum).

3) To demonstrate that disadvantaged Israeli children can do well in such a program, exceeding the level of functioning manifested in their regular school activities.

4) To compare the predictive validity of IQ and LP in predicting standardized school achievement of Israeli children.

5) To compare the predictive validity of IQ and LP in predicting achievement of Israeli children in the "Letters and Numbers" program.

6) To compare the relative validity of LP and IQ (as in objectives 4 and 5) separately for three different groups of home background.

7) To identify the factor structure in a field containing verbal and nonverbal intelligence, LP, standardized school achievement, and "Letters and Numbers" performance.

8) To compare the factor structures for the three different groups of home background.

Chapter 3: Method

Subjects: The subjects were selected from a large representative sample of the Israeli elementary school children (used in a large evaluation study of the education of disadvantaged children in Israel directed by Minkovitch, Davis, and Bashi and supported by the Van Leer Foundation and the Israeli Ministry of Education and culture). In the framework of that large evaluation study, a comprehensive battery of instruments was administered in May, 1973. The present sample consists of children who were tested in the sixth grade. The letters and Numbers program was administered in Spring, 1974, while the subjects were in the seventh grade.

Five Jerusalem schools were selected, two defined as schools for disadvantaged children and the other three representing the entire range of home backgrounds. One school was dropped from the sample due to faulty administration of tests, so that the final sample consisted of 8 classes in 4 schools. In total, 200 children participated in the Letters and Numbers program, 45% of which were boys.

Background, was defined by the maximal education of the father or the mother, and was distributed as follows: 33.8% of the subjects' parents had 7 or less years of education, 49.4% of the subjects' parents had between 8 and 11 years of schooling, and 16.9% of the subjects' parents had 12 or more years of schooling.

The mean IQ of the tested sample on the "MILTA" verbal intelligence test (Ortar, 1965) was 100.01, the standard deviation was 12.1 points and the IQ scores ranged from 71 to 133.

The tests were administered by trained students, under the supervision of the principal investigators. Experienced teachers (on their sabbatical year) were chosen to teach the program in the schools. The sessions of the Letters and Numbers program were held in the home classrooms, and the students were not divided according to the groupings in arithmetic.

Materials

I. The LP tests: 1) the SLPT. The series Learning Potential Test (Babad and Budoff, 1971) is a group test for the elementary grades. The task in the SLPT is the completion of series of pictures or geometric forms, arranged in a pattern in which the figures change systematically. Each item presents a horizontal row of cells, containing a stimulus figure. One cell is left blank. The subjects must identify among the multiple choices the picture or form which best completes the series. Four concepts may vary in a series: semantic content (meaningful or geometric figures), size (large/small), color (black/white), or orientation (up/down or left/right). The concepts may vary symmetrically or asymmetrically and the blank space may be placed in any part of the series.

The version of the SLPT used in the project consisted of two equivalent 50-item forms. (Of the 50 test items, 30 were picture series items, and 20 were geometric forms series, ranging in difficulty.) The coaching booklet contained 20 picture series items, none identical to any of the test items.

The SLPT was administered / in three regular (45-minute) class sessions, within a period of 3-5 days. In the second session, group coaching was provided to the entire class. In that session, the following strategies were taught and practiced by the entire class in a relaxed, non-competitive atmosphere:

1)The concept of a series can be conceived as a "tune" with a distinct rhythm. Forming the concept is practiced via "singing the tunes".

2) Most series problems combine more than one concept (tunes along the color, size, object or position dimensions).

For effective solution, one must separate the tunes, and find the options correct for that tune regardless of the other dimensions. The final solution is reached by elimination for each concept.

3) One should search for the starting point of each tune, keeping in mind that the starting points do not always correspond to the first stimulus in the item.

4) Tunes can be sung forward and backward, the decision depending on the location of the blank cell.

2.) The MLPT. The version developed for the present study consisted of equivalent forms containing 43 double classification items. Of these items, 28 were Raven's Progressive Matrices and the rest were constructed for this test. Most of the items (33) were 2 x 2 classifications and the rest 3x3 classifications. The coahing booklet consisted of 14 double classification items, none of which appears in the test, but similar to the test items in the reasoning process required for solution.

Like the SLPT, the MLPT was also administered in three regular class sessions, in a "pretest-coach-posttest" sequence. Training was conceptually similar to the SLPT training, emphasizing the separate solution for every element, the flexibility in looking at the problems, and the reduction of memory load by step-by-step elimination.

(black/white), or orientation (up/down or left/right). The concepts may vary symmetrically or asymmetrically and the blank space may be placed in any part of the series.

The version of the SLPT used in the project consisted of two equivalent 50-item forms. (Of the 50 test items, 30 were picture series items, and 20 were geometric forms series, ranging in difficulty.) The coaching booklet contained 20 picture series items, none identical to any of the test items.

The SLPT was administered / in three regular (45-minute) class sessions, within a period of 3-5 days. In the second session, group coaching was provided to the entire class. In that session, the following strategies were taught and practiced by the entire class in a relaxed, non-competitive atmosphere:

1)The concept of a series can be conceived as a "tune" with a distinct rhythm. Forming the concept is practiced via "singing the tunes".

2) Most series problems combine more than one concept (tunes along the color, size, object or position dimensions).

For effective solution, one must separate the tunes, and find the options correct for that tune regardless of the other dimensions. The final solution is reached by elimination for each concept.

3) One should search for the starting point of each tune, keeping in mind that the starting points do not always correspond to the first stimulus in the item.

4) Tunes can be sung forward and backward, the decision depending on the location of the blank cell.

2) The MLPT. The version developed for the present study consisted of equivalent forms containing 43 double classification items. Of these items, 28 were Raven's Progressive Matrices and the rest were constructed for this test. Most of the items (33) were 2 x 2 classifications and the rest 3x3 classifications. The coahing booklet consisted of 14 double classification items, none of which appears in the test, but similar to the test items in the reasoning process required for solution.

Like the SLPT, the MLPT was also administered in three regular class sessions, in a "pretest-coach-posttest" sequence. Training was conceptually similar to the SLPT training, emphasizing the separate solution for every element, the flexibility in looking at the problems, and the reduction of memory load by step-by-step elimination.

The following scores were derived from the LP tests:

1) The SLPT pretest score.
2) The MLPT pretest score.
3) The LP total pretest score. This score is the sum of the two previous scores.
4) The SLPT posttest score.
5) The MLPT posttest score.
6) The LP total posttest score, which is the sum of the two previous scores.

The LP total scores are combined measures created so that a singular index could be used in analyses, an index which is more reliable (and hopefully more valid) than it's components. The IQ score, to which LP is compared in this study, is also a composite score, but unlike the LP score -- the components of which are conceptually homogeneous - the IQ score combines a variety of conceptually distinct subtests.

II. The IQ test. Verbal IQ was measured by the "MILTA" (Ortar, 1965), the only standardized group verbal intelligence test for this age group in Israel. The test consists of four subtests: Vocabulary, Sentence Completion, Odd-out, and Arithmetic Problems. Raw scores for each subtest are transformed into scaled scores, and the sum of scaled scores is converted into one IQ score. The test is administered in one regular class session.

III. Level I test. A substitution test was used to measure Level I ability. The test was taken from "MILKAN" (Ortar and Babad, 1973). Six pictures are used in the test, and the number of items is 144. Administration of the test takes about 7-8 minutes, of which 150 seconds is the actual time of the test.

IV. Achievement tests. 1) Arithmetic. A standardized test was used consisting of 37 multiple choice items (operations and problems). The test is based on the requirements of the sixth grade curriculum.

2)Reading. A standardized achievement test was used, consisting of 54 multiple choice reading comprehension items. Both achievement tests were administered in one regular class session.

V. The Letters and Numbers test. The criterion in this study was the cryptarithmetic performance. The test represented all parts of the Letters and Numbers program. It consisted of 24 items, but due to printing errors, only 23 items were scored. Four equivalent forms of the test were used. The test was administered in one class session.

Procedure

The intelligence test, the Level I test, the two achievement tests and the background data were collected in May 1973, as part of the evaluation study of the education of disadvantaged children in Israel. All tests were administered within one week, when the subjects of the present study were in the sixth grade. The LP tests were administered in March, 1974, when the subjects were in the seventh grade. Care was taken that both LP pretests would be administered before any coaching was given. The coaching and posttesting of the SLPT preceeded that of the MLPT.

The Letters and Numbers program was taught from April until June 1974. The program was scheduled for 15 sessions, in which the students completed the 25 work sheets. The teachers distributed the work sheets one at the time, keeping the entire class roughly in the same part of the program. Faster students were encouraged to help slower students. Frontal teaching was minimized but not disallowed, and the teachers tutored and supervised the students' work, providing additional support when necessary.

Additional sessions for slower students were provided when deemed necessary by the teachers. To prevent the formation of self fulfilling exepectancies. Our teachers were not given any information about the academic standing of any student. The teachers' perceived their task as bringing all students to the criterion level, by creating a multiple support system in the classroom.

Bi-weekly supervision meetings were held with the teachers by the principal investigators. In these meetings, the teachers reported the progress of the program, compared notes and discussed special problems and difficulties that they encountered. These meetings resulted in continuous modifications of the program.

The teachers reported that the students were highly interested in the program, and that their interest and motivation were maintained throughout the period. Some of the students started inventing cryptarithmetic problems, while other tried to solve the problems appearing in the weekend papers.

Chapter 4: Results and discussion.

Means and distributions.

The means and standard deviations of the different variables are presented in table 1.

Table 1. Means and Standard deviation of the major variables.

Variables		Number of items in test.	Mean	SD	N
LP scores	MLPT pretest	43	31.46	7.29	197
	MLPT posttest	43	34.12	5.79	197
	SLPT pretest	50	24.45	7.94	180
	SLPT posttest	50	36.24	5.63	180
	LP pretest (total)	93	55.91	12.69	180
	LP posttest (total)	93	70.36	9.92	180
Standardized Intelligence.	Level I(substitution)	44	77.94	40.11	170
	Verbal IQ	Standardized	100.01	12.10	169
	Math. in IQ test	Standardized	10.27	2.39	169
Standardized Achievement	Reading	54	27.72	10.83	90
	Math	37	14.81	5.37	134
Letters and Numbers	Criterion	23	15.83	5.38	184

As can be seen, the improvement from pretest to posttest was quite large for the

SLPT, and somewhat smaller for the MLPT. Our hypothesis is that individual learning and improvement takes place already in the MLPT pretest, while in the SLPT the students depend more on the strategies taught in the coaching session. Thus the posttest-pretest difference is larger for the SLPT. We shall elaborate on this issue in our discussion of the correlation matrix.

The LP posttest mean is much higher than the LP pretest mean(d= 14.45, which is 1.14 SD). The variances of the LP tests deacreased from pretest to posttest by 20-30%. This finding is in line with Babad and Budoff's (1974) argument that numerous false positives change their relative positions in the distribution, thus eliminating the spread typically found for disadvantaged children in the lower part of the distribution resulting in smaller variation.

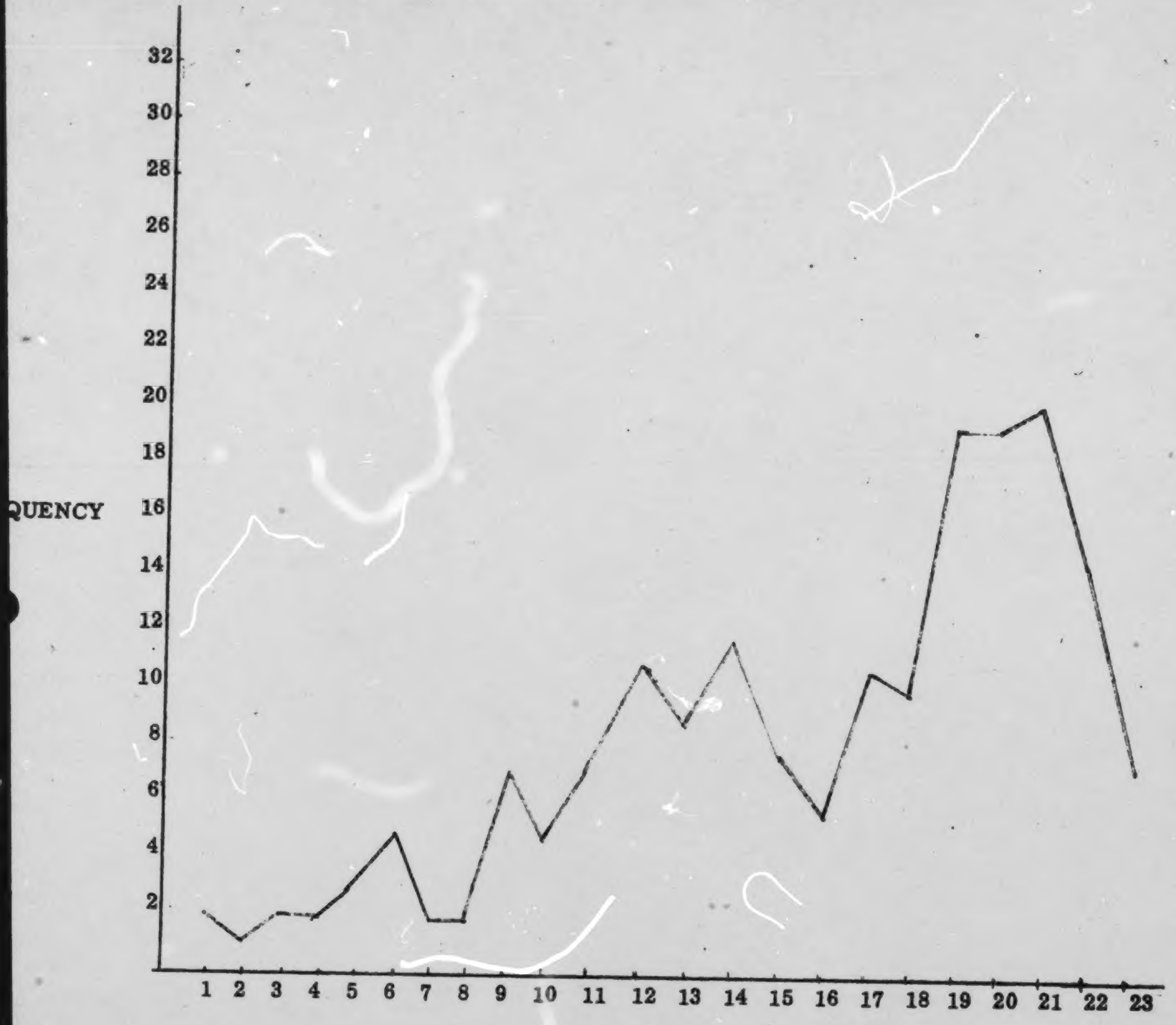
The mean of the verbal IQ in the sample (100.01) and the mean of it's math component (10.27), match the standardized means of the israeli population. Since no retarded and no exceptionally gifted children were included in the sample, the variances, as expected, are smaller than in the total population.

Data were initially computed separately for boys and girls, searching for sex differences. No such differences were found, and except for some random fluctuations the means for boys and girls were almost identical. Therefore, we did not present separate results for boys and girls.

The distribution of Letters and Numbers scores is presented in Figure 1. The mean of this distribution (in Table 1) is 15.83 (out of 23 possible points), and the standard deviation is 5.38. As can be seen in figure 1, the distribution is skewed toward a ceiling effect, indicating that most children have indeed attained mastery, and they can competently solve cryptarithmetic problems (The ceiling effect is indicated by the fact that although the mean of the distribution is 15.83, the mode is 21 and the highest frequencies are found for the scores 19-22.

Only 10% of the subjects were incapable of solving a third of the problems while 54.3% solved at least two thirds of the problems. The fact that the distribution is skewed toward a ceiling effect and yet the variance is large enough is extremely interesting. The program was designed to provide an atmosphere and a learning method which would facilitate the attainment of mastery by all children. Progress from one task to the next one was determined by the attainment of mastery by the children (including compensatory sessions with the slower learners). Since this approach aims at general mastery, chances of reaching a ceiling effect are very high, and therefore the variance is expected to be very low. Such a situation is quite satisfactory for a designer of a modern curriculum, but it makes empirical validity research almost impossible, unless other criteria, such as time to reach mastery, are used to validate the predictors. The results in figure 1 (a tendency toward a ceiling effect and a large enough variance) are satisfactory from both points of view. However, since the norm-referenced and the criterion-reffrenced approaches almost represent a contradiction in terms, the conceptual-methodological problem here merits special consideration beyond the scope of this study.

Figure 1. Distribution of scores in the Letters and Numbers test.



SCORE IN LETTERS AND NUMBERS TEST.

Correlations matrices.

In table 2 we present the correlation matrix between all test scores (N=200). Some of the correlations (in the LP and IQ sections) are whole-part correlations, and these are underlined.

In the LP-LP quadrant, one can see that the correlation between pretest and posttest MLPT is high (.75), but rather low (.26) for the SLPT. This difference is probably due to the fact the MLPT is very graded, and includes enough items of a given type. This provides repetition of the same mental process which constitutes training. The self learning is enhanced because some of the students are able to develop the required strategies without external intervention. The case is different in the SLPT, where the most effective strategy ("tunes") is not readily available and usually is created by explicit instruction. Therefore correlation between pretest and posttest SLPT is rather low, but the improvement following training (in Table 1) is large.

The correlation between the two LP pretests is .38, while the correlation between the two LP posttest is .50. In the pretests the two tasks are different from each other (in spite of the fact that both are concept fromation tests) and the common variance is rather small. In the posttests, where both scores include the effects of strategy learning, the amount of common variance is larger.

Table 2: Pearson Correlations between test scores for the total sample (N=200)

		LP					Standardized Intelligence			Standardized Achievement		Letters & Numbers
		2 MLPT Posttest	3 SLPT Pretest	4 SLPT Posttest	5 LP total Pretest	6 LP total Posttest	7 Level I	8 Verbal IQ	9 Math. in Verbal IQ	10 Reading	11 Math	12 Criterion
LP	1. MLPT Pretest	.75	.38	.51	.81*	.73	.03	.42	.27	.45	.44	.59
	2. MLPT Posttest		.33	.50	.63	.87	.12	.44	.36	.38	.43	.56
	3. SLPT Pretest			.26	.85	.33	.08	.32	.26	.05	.30	.22
	4. SLPT Posttest				.47	.86	.10	.45	.35	.40	.51	.60
	5. LP total Pretest					.63	.04	.44	.32	.26	.44	.47
	6. LP total Posttest						.13	.51	.39	.46	.54	.66
Standardized Intelligence	7. Level I							.09	.02	.15	.00	.02
	8. Verbal IQ								.73	.47	.54	.52
	9. Math. Subtest in Verbal IQ									.37	.53	.47
Standardized Achievement	10. Reading										.34	.46
	11. Math											.59

* Part-whole relationships are underlined.

The LP total posttest score is almost identical to the verbal IQ score in predicting standardized achievement in reading and math. This finding (with other findings quoted in chapter 1) is quite suprising if one notes that the IQ test in this study is composed of a verbal component and a math component while the LP tests are non-verbal and non-arithmetical.

As expected, the Level I test (substitution) is unrelated to the other variables under investigation. In the next analyses, this variable will be omitted.

The score in the Letters and Numbers criterion test is best predicted from the LP total posttest score. The MLPT predicts as well in pretest and in posttest, while in the SLPT the improvement in prediction from pretest to posttest is large. The LP total posttest score is a more powerful predictor than its two components. The verbal IQ score is inferior to the LP total posttest score in predicting the Letters and Numbers criterion. Interestingly, even the math component of the IQ test is a less effective predictor than the total verbal IQ, in spite of it's contentual similarity to the criterion. The power of the standardized achievement score in math to predict the criterion was somewhere between the verbal IQ score and the LP total posttest score.

In conclusion, the tests with the best chances of predicting success in the Letters and Numbers program are those which provide learning pertaining to conceptual (Level II) ability, while the IQ tests and the standardized achievement tests (reflecting conventional school learning) are less effective predictors.

Results of several analysis of eighth-grade surveys in Israel (Smilansky and Yam, 1967, Ortar, 1967, and Davis 1974) indicate that of all background variables, the most crucial is parents' education.
In order to examine the pattern of relationships and the validity of LP and IQ in each group separately, we categorically divided the parents' maximal education into three groups ("low background" - less than 8 years of schooling (33.8% of the sample); "middle background" - between 8 and 11 years of schooling (49.4% of the sample); and "high background" 12 and above years of schooling (16.9% of the sample).

The first impressive finding is that the three background groups do not differ from each other in their performance on the Letters and Numbers criterion test.

The means are 16.18, 16.34 and 17.56 for low, middle, and high background groups, respectively. A one-way analysis of variance yields no significant difference between the groups. This finding indicates that when learning is started from an equalized starting point and is carried on while giving enough opportunity for steady progress and when the creation of any cumulative deficit is avoided, groups differing in educational background, can reach the same level of conceptual (Level II) thinking (at least on cryptarithmetic problems).

In Tables 3, 4, and 5 we present the separate correlation matrices for each background group. Without exception, the LP total posttest score is the best predictor of the Letters and Numbers criterion. The superior validity of that variable is particularily salient in the low background group, which is the target population of the LP tests.

Typical school achievement (as measured by the standardized tests) and achievement in this specially designed program (Letters and Numbers test) are two types of criteria for examining the relative validity of LP and IQ across the educational background groups. Regarding standardized achievements, it seems from the tables that IQ is most effective as a predictor in the high background group (Table 5, correlations .62 and .93), and less effective in the lower two groups (Tables 3 and 4, range of correlations .45 to .64). The LP total posttest score, on the other hand, is most effective as a predictor in the low background group (Table 3, correlations .53 and .69), and less effective in the higher two groups (Tables 4 and 5, range of correlations .37 to .51). The separate comparisons between LP and IQ in each table show that LP is superior to IQ in predicting standardized achievement in the low background group (Table 3), more or less similar to IQ in the middle background group (Table 4), and clearly inferior to IQ in the high background group (Table 5). Thus, the two types of tests show reversed trends in their effectiveness as predictors of standardized achievement: IQ improving with higher background, LP improving with lower background. As we claimed earlier, the IQ test is middle-class oriented, based on assumptions that are more fully realized for children growing up in a rich environment, while the LP test is especially designed for children who are environmentally deprived. It seems that the LP posttest is a good facilitator of the manifestation of hidden potential of disadvantaged

Table 3: Correlations between test scores for children of low educational background (N=50)

		LP					Standardized Intelligence		Standardized achievement		Letters & Numbers
		2 MLPT posttest	3 SLPT pretest	4 SLPT posttest	5 LP total pretest	6 LP total posttest	7 Verbal IQ	8 Math in Verbal IQ	9 Reading	10 Math	11 Criterion
LP	1. MLPT pretest	.66	.48	.49	.86*	.67	.38	.17	.17	.45	.59
	2. MLPT posttest		.36	.47	.59	.90	.42	.28	.48	.60	.70
	3. SLPT pretest			.21	.86	.32	.33	.30	.07	.24	.22
	4. SLPT posttest				.41	.81	.63	.51	.40	.57	.70
	5. LP total Pretest					.58	.41	.27	.26	.41	.47
	6. LP total Posttest						.59	.43	.53	.69	.79
Standardized Intelligence	7. Verbal IQ							.79	.45	.64	.54
	8. Math subtest in Verbal IQ								.37	.54	.60
Standardized Achievement	9. Reading									.48	.62
	10. Math										.62

* Part-whole relationships are underlined.

Table 4. Correlations between test scores for children of middle educational background (N=75)

		L P					Standardized Intelligence		Standardized Achievement		Letters & Numbers
		2 MLPT posttest	3 SLPT pretest	4 SLPT posttest	5 LP total pretest	6 LP total posttest	7 Verbal IQ	8 Math in Verbal IQ	9 Reading	10 Math	11 Criterion
L P	1. MLPT pretest	.79	.33	.54	.75*	.77	.46	.35	.55	.46	.59
	2. MLPT posttest		.25	.55	.59	.86	.46	.43	.27	.44	.57
	3. SLPT pretest			.23	.88	.24	.35	.29	.05	.34	.24
	4. SLPT posttest				.44	.81	.40	.20	.36	.47	.55
	5. LP total Pretest					.58	.49	.39	.33	.48	.48
	6. LP total Posttest						.48	.32	.46	.51	.64
Standardized Intelligence	7. Verbal IQ							.65	.46	.47	.44
	8. Math subtest in Verbal IQ								.27	.48	.43
Standardized achievement	9. Reading									.28	.34
	10. Math										.58

* Part-whole relationships are underlined.

-26 b-

Table 5: Correlations between test scores for children of high educational background (N=26)

		LP					Standardized intelligence		Standardized Achievement		Letters & Numbers
		2 MLPT posttest	3 SLPT pretest	4 SLPT posttest	5 LP total pretest	6 LP total posttest	7 Verbal IQ	8 Math in Verbal IQ	9 Reading	10 Math	11 Criterion
L P	1. MLPT pretest	.60	.34	.62	.85 *	.77	.17	.08	.25	.21	.50
	2. MLPT posttest		.52	.70	.82	.92	.39	.23	.44	.15	.55
	3. SLPT pretest			.43	.78	.51	.39	.01	.43	.40	.34
	4. SLPT posttest				.65	.93	.49	.42	.49	.50	.67
	5. LP total Pretest					.79	.33	.06	.38	.40	.54
	6. LP total Posttest						.48	.36	.50	.37	.67
Standardized intelligence	7. Verbal IQ							.71	.93	.62	.52
	8. Math subtest in verbal IQ								.66	.71	.37
Standardized achievement	9. Reading									.27	.41
	10. Math										.40

* Part -whole relationships are underlined.

children which the IQ test fails to identify.

The picture is quite different when we compare the validity of LP and IQ in predicting the Letters and Numbers criterion score. Here the LP total posttest score is superior to the verbal IQ in all three background groups (.79 vs. .57 in table 3; 64 vs. .44 in Table 4; and .67 vs. .52 in Table 5). The difference between the predictions of LP and IQ is the largest in the low background group, and it decreases with improved educational background. The LP total posttest score is also superior to both standardized achievement test scores in predicting the criterion.

Conceptually, comparing the validity of LP, IQ and standardized achievement in predicting Letters and Numbers performance means confronting the content and process dimensions. The content dimension consists of verbal and mathematical elements emphasized both in the IQ test and in the standardized achievement tests. The process dimension, emphasized in the LP tests and supposedly central in school work as well, involves the learning of relevant problem solving strategies and applying them to specific problems following ample training. The Letters and Numbers criterion has the content in common with the IQ and the standardized achievement tests, and the process in common with the LP tests. We found that in all background groups LP was a superior predictor to all other variables although it is the remotest from the criterion on the content dimension. Even the standardized math achievement test - the most similar in content to the criterion - was less effective than LP.

It seems that the superiority of LP in predicting the criterion is due to some central processes common to both. These processes are the basis of LP, not only as a method of mental testing, but also as a conceptual framework for educating disadvantaged children: All learners enter a program at an equalized starting point which poses minimal demands, and all of them master it's prerequisites; all learners are provided with sufficient opportunity to reach mastery at any stage in a supportive system so that no deficit is accumulated; bits of information presented at any stage are regulated to accomodate the pace of all learners; the program is process-oriented, requiring no memorization and stressing the learning of strategies and their application to problems graded in difficulty level.

When a new, well designed learning program is based on no previous deficits, the low background group has better chances of success even if the methods and the contents of the program are previously unknown to children. If indeed methods, contents and processes of school learning would advance in a proper pace from equalized starting points, the chances of minimizing the cultural gap look somewhat better.

Factor analyses.

The factor analysis on the total sample (Table 6) yielded 3 factors, accounting for 67% of the variance. As can be seen in Table 6, the first factor, accounting for 45.6% of the variance, is highly loaded on the verbal IQ (.98) and on it's verbal and math components. The standardized achievements (.50 and .56) are also included in this factor. The Letters and Numbers criterion is relatively highly loaded on the first (.56) and on the third (.46) factor. The first factor is labeled "Verbal and school ability", reflecting it's verbal component and it's relationship with actual school work.

The second factor, accounting for 11.3% of the variance, is highly loaded on the LP pretests (.91). We interpret this factor to reflect the subjects' "one-shot" tested non-verbal intelligence, excluding effects of learning. The factor label "Non-verbal intelligence" refers to Jensen's Level II ability.

The third factor, accounting for 10.1% of the variance, is highly loaded on the LP posttests (.76) and also on the MLPT pretest (.58) and the Letters and Numbers criterion (.46). We interpret this factor to represent the students' "Potential for strategy learning", distinct from their ability to solve non-verbal problems (on pretests) which constitute Factor 2.

The factor analysis on the total sample leads to the following conclusions:

1) The LP posttests are clearly distinguished from the LP pretests.

2) The LP tests are distinguished from the verbal IQ tests.

3) The achievement tests are highly loaded on the verbal factor.

4) The criterion is highly loaded (and almost evenly divided) on the verbal factor and the strategy learning factor.

These findings support the demand for using a "test-coach-test" paradigm. They also provide further confirmation for the argument that the criterion has a content dimension in common with IQ and standardized achievement tests, and a process dimension in common with the LP posttest performance.

Since the correlation patterns of the three background groups were distinct from each other, we decided to repeat the factor analysis separately for groups of low, middle, and high educational background, in order to examine the three factor structures. The results are shown in Table 6. The three separate analyses have much in common with the analysis of the total sample. The range of variance explained in these analyses is 66%-77%. The first factor is rather large in all three analyses, accounting for 45%-50%. The third factor in these analyses accounts for 9%-11% of the variance.

The first factor in all three analyses has an extremely high loading on one variable (.94 and higher): verbal IQ in the high background group, LP total posttest score in the middle background group, and the LP total pretest score in the low background group. The "Non-Verbal Intelligence" factor accounted for almost 50% of the variance in the low background group but only for 10% or less of the variance in the other background groups. The "Verbal and school ability" factor accounted for 50% of the variance in the high background group, while accounting for less than 15% of variance in the two lower background groups. The "Potential for strategy learning" factor was central in the middle background group (45% of the variance), but more peripheral in the high background group (18% of the variance) and didn't exist as a unique factor in the low background group. Thus, the results indicate that the factor structures in the three background groups are indeed quite distinct of each other, not so much in what the factors represent, but in their relative orders of magnitude.

An interesting pattern is found for the standardized achievement scores. In the high background group, these scores are highly loaded on the verbal factor which is the first factor in that analysis. In the other two groups the loadings of the achievement scores are more spread among the factors, and are not so highly loaded on the verbal factor. In the middle background group the loadings of the achievement scores on the first factor are lower than in the high background group, and in the low background group their loading on the first factor are very low.

Table 6. Factor analyses for total sample and for the three background groups.

Variables		Total sample (N=200)			Low background (N=50)			Middle background (N=75)			High background (N=25)		
		Factor 1 (45.6% of variance)	Factor 2 (11.3% of variance)	Factor 3 (10.1% of variance)	Factor 1 (48% of variance)	Factor 2 (14% of variance)	Factor 3 (11% of variance)	Factor 1 (45% of variance)	Factor 2 (11% of variance)	Factor 3 (10% of variance)	Factor 1 (50% of variance)	Factor 2 (18% of variance	Factor 3 (9% of variance)
L P	MLPT pretest	.34	.53	.58	.82	.20	.21	.76	.22	.34	.04	.77	.41
	MLPT posttest	.39	.40	.61	.67	.33	.39	.75	.20	.28	.22	.68	.52
	SLPT pretest	.15	.88	-.09	.72	.17	-.45	.03	.12	.94	.25	.13	.76
	SLPT posttest	.47	.21	.58	.36	.60	.33	.75	.17	.13	.44	.72	.26
	LP total pretest	.27	.91	.28	.97	.19	-.18	.43	.19	.86	.14	.59	.79
	LP total posttest	.49	.34	.76	.63	.53	.47	.94	.19	.18	.37	.80	.42
Standardized Intelligence	Level I (substitution	-.06	-.10	.25	-.07	.14	.64	-.04	-.31	.03	.08	.66	.02
	Verbal IQ	.98	.20	-.07	.19	.99	-.07	.25	.94	.25	.97	.05	.23
	Verbal component of IQ (3 subtests)	.87	.20	-.01	.23	.83	-.01	.25	.84	.18	.90	.06	.41
	Math component of IQ (1 subtest)	.71	.14	.02	.11	.78	-.03	.22	.51	.27	.84	.21	-.27
Standardized Achievement	Reading	.50	.02	.34	.25	.42	.49	.43	.36	.02	.80	.06	.32
	Math	.56	.24	.27	.37	.61	.24	.47	.32	.29	.61	.20	.09
Letters and Numbers	Criterion	.56	.21	.46	.46	.57	.18	.61	.30	.21	.43	.46	.27
Factor Label:		Verbal and school ability	Non-verbal Intelligence	Potential for Strategy Learning	Non-verbal Intelligence	Verbal and school ability	skill and Practice	Potential for Strategy and School Learning	Verbal Intelligence	Non-verbal intelligence	Verbal and school ability	Potential for Strategy Learning	Non-verbal Intelligence

The letters and Numbers criterion is clearly loaded on one facaor only for the middle background group ("Potential for strategy learning" factor). The criterion's loadings are evenly divided between the three factors for the low background group and evenly divided between two factors for the high background group.

The results confirm the LP conceptualization, showing a clear factorial distinction between the pretest and the posttest performance. The posttest performance reflects cognitive learning that is manifested in neither the IQ test nor the LP pretest. Thus, the validity of the "training-within-test" paradigm is confirmed by the findings.

Multiple regression analysis.

In order to determine the relative validity of each major variable, controlling for it's relationships with all other predictors, we carried out a multiple regression analysis using the stepwise technique. The letters and Numbers score served as the criterion in this analysis. Since the specific computer program used, rejects any subject with missing data, the number of subjects in this analysis was rather small (N=69). It was therefore impossible to carry out separate analyses for the three background groups. Therefore, background was entered as a predictor in this analysis, using the categories (low, middle, high) as the three values.

In the multiple regression analysis presented in Table 7, no order of entry was forced. The magnitude of prediction of each variable and it's level of significance are presented in the first two columns of Table 10. The "R square" column provides the accumulated proportion of explained variance following each step, and the "R square change" column lists the proportion of criterion variance explained by each variable separately.

As can be seen in Table 7, all predictors together accounted for 54% of the criterion variance. The LP total posttest score was the major predictor, accounting for 43% of the variance. The second most significant predictor was the standardized math score, which accounted for 8% of the variance. The standardized reading score was marginally significant as a predictor, while the IQ and the LP total pretest scores were insignificant predictors, accounting for 1% of the variance or less.

Table 7. Multiple regression analysis using stepwise regression technique, carried on the total sample (N=69)

Variables entered: Background (3 groups), LP total pretest, LP total posttest, Verbal IQ, Standardized reading achievement, Standardized math achievement.

Criterion: The Letters and Numbers score.

Step	Variable	F value	Significance	R square	R square change (% of variance)
1	LP total posttest	50.55	$p < .001$	.43	.43
2	Standardized math achievement	10.67	$p < .002$	.51	.08
3	Standardized reading achievement	2.84	$p < .10$	.53	.02
4	Verbal IQ	< 1	n.s.	.54	.01
5	LP total pretest	< 1	n.s.	.54	.00

Total variance explained: 54%

Background thrown out of regression analysis due to it's low correlation (.05) with the criterion.

In spite of the clear results in the hypothesized direction, we felt that additional examinations must be carried out in order to clarify the status of the verbal IQ score. Two additional multiple regression analyses were computed, forcing the order of entry of verbal IQ and the LP posttest variables. In the first analysis, we determined that the LP posttest score would enter first, the verbal IQ score would enter second, and the other variables would then follow with no determined order. In the second analysis we reversed the order; verbal IQ entered first, the LP posttest score entered second, and the other variables followed in an undetermined order. The findings of these two additional multiple regression analyses can be summarized as follows:

1) Verbal IQ was found a significant predictor ($p < .05$) in both analyses.

2) The standardized math score, forced into the third place, remained a significant predictor ($p < .05$), accounting for 5% of the variance in both analyses.

3) When the LP posttest score proceeded the verbal IQ, it still accounted for 21% of the variance, additional to 27% accounted for by the verbal IQ. When the verbal IQ proceeded the LP posttest score, it added only 4% to the 43% accounted for by the LP variable.

The comparative results of the three multiple regression analyses re-confirm the validity of the LP posttest score in predicting attainment in the Letters and Numbers program, in clear superiority over any other predictor used in the study. The LP tests, designed to identify potentially able children among the disadvantaged, were found the most valid predictors of Level II functioning in a learning program designed to minimize the penalty paid by culturally deprived children for their accumulated deficits.

REFERENCES

Babad, E. Y., and Bashi, J. Unpublished data. School of education, Hebrew University of Jerusalem, 1975.

Babad, E. Y., & Budoff, M. A preliminary manual of the Series Learning Potential Test. Unpublished manuscrip, Research Institute for Educational Problems, 1971.

Babad, E. Y., & Budoff, M. Sensitivity and validity of learning-potential measurement in three levels of ability. *Journal of Educational Psychology*, 1974, *66*, 439-447.

Budoff, M. Learning potential: A supplementary procedure for assessing the ability to reason. Seminars in Psychiatry, 1969, I, 278-290.

Budoff, M., Corman, L., & Gimon A. Educational test of learning potential assessment with Spanish speaking youth. *Studies in Learning Potential*, 1974, *4*, No. 71.

Budoff, M., Corman, L., & Litzinger, C. S. A Manual for the Raven Learning Potential Test. *Studies in Learning Potential*, 1973, *3*, No. 59.

Budoff, M., Gimon, A., & Corman, L. Learning potential measurement with Spanish-speaking youth as an alternative to IQ tests: a first report. *Studies in Learning Potential*, 1974, *4*, No. 70.

Budoff, M., Meskin, J., & Harrison, R. H. An educational test of the learning potential hypothesis. *American Journal of Mental Deficiency*, 1971, *76*, 159-169.

Corman, L., & Budoff, M. The Series test as a measure of learning potential. *Studies in Learning Potential*, 1973, *3*, No. 47.

Davis, D. Unpublished data. School of Education, Hebrew University of Jerusalem, 1974.

Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., & Tyler, R. *Intelligence and cultural differences.* Chicago: University of Chicago Press, 1951.

Feuerstein, R. Learning potential assessment device. In B. W. Richards (Ed.). *Proceedings of the First Congress of the International Association for the Scientific Study of Mental Deficiency,* Reigate, Surrey: Michael Jackson, 1968.

Flanagan, J. C. *Test of General Ability: Technical report*. Chicago: Science Research Associates, 1960.

Hunt, J. McV. Intelligence and experience. New York: Ronald Press, 1961.

Jacobs, Paul I. and Mary Vandeventer: The learning and transfer of Double - Classification Skills by first graders
in Child development. March 1971. Vol. 42, no. 1, pp. 149-159 .

Jensen, A. R. How much can we boost IQ and scholastic achievement ? Harvard Educational Review, 1969, 39, 1-123.

Newell, A. & Simon, H. A. Human problem solving. Englewood Cliffs, N. J. : Prentice Hall, 1972.

Ortar, G. Improving test validity by coaching. Educational Research, 1960, 2, 137-142.

Ortar, G. Milta, School of Education, Hebrew University of Jerusalem, 1965.

Ortar, G. Thirteen Years of "Seker". Megamot, 1967, 15. (In Hebrew)

Ortar, G., & Babad, E. Y. Milkan, School of Education, Hebrew University of Jerusalem, 1973.

Smilansky, M., & Yam, Y. The relations between size of family, father's education, cognitive abilities, and school achievements. Megamot, 1967, 16 (in Hebrew)

Zigler, E. Mental retardation: Current issues and approaches. In L. W. Hoffman & M. L. Hoffman (Eds.), Review of child development research. Vol. 2. New York: Russell Sage Foundation, 1966.

Part III. Summary and Implications 1 - 8

Summary and Implications

The data we have accumulated continues to indicate a broad range of consistency and support for the differences shown by psychometrically defined EMRs in response to the test-train-retest paradigm which constitutes the learning potential assessment strategy. More able children, by the Learning Potential criterion, indicate areas of more adequate functioning in aspects of a child's life that do not depend on satisfactory progress in school. In this project, those children who have demonstrated their competence on the LP task also score higher on other nonverbal or minimally verbal tasks, e.g., WISC, Performance IQ, and Raven's Progressive Matrices. The personality data, whether behavioral or verbal, suggests a greater sense of personal adequacy, and a greater sense of realism in their vocational planning.

Competence and understanding, more generally, can be demonstrated in the more able (LP) child's ability to build empirically meaningful relationships among objects, and/or events, without being required to either abstract common elements or to verbalize one's empirical understanding. The more able (LP) students' ability to perform in this manner was clearly demonstrated in the study in which the child was taught principles of electricity in which he could learn by manipulating batteries, bulbs, and wires. Following teaching, they were able to display understanding of principles of electricity to the same degree as taught low achieving, higher IQ children, and in some instances, as well as average achieving children. Their salient failure was their inability to verbalize causal reasons for their choice. But there are probably large numbers of adults who function very adequately occupationally, socially, and interpersonnally, without having to resort to verbal conceptual explanations of their behavior. They empirically understand, and they respond appropriately to their environment.

Clearly, then, if one wished to define ability in the broader sense of seeking a comprehensive definition of capability, especially when the child is not readily able to perform on the common verbally-biased tests used, it is incumbent on psychologists to search for parameters that are minimally associated with verbal facility and that are related to a diverse range of nonschool behaviors that are relevant to the person's prior and/or current experiences as the critical indicators of ability to learn and profit from experience.

Traditional testing formats have failed to develop and/or make use of this principle as one which must govern judgments of capacity for future learning and/or cognitive growth. It is this factor which has resulted in the high incidence of misclassification among children from social backgrounds other than the white, middle income, English-speaking, urban area dwelling children who perform most satisfactorily on tests and in schools.

This project has provided further support for the prior findings by demonstrating the utility of another learning potential instrument constructed around training appropriate to the Raven Progressive Matrices, a nonverbal abstract reasoning task. Further, there is little loss in the use of the procedure within a small group format. With an IQ-identified EMR population, the groups should be smaller (probably not over six children with an experienced trainer well versed in the training procedure). Most interesting is the finding reported from several sources, that posttraining scores account for significant sources of variance with clearly educational tasks. In this project, posttraining scores but not IQ accounted for significant unique proportions of the variance related to the vocational-social and even, the academic curricula. The latter finding was unexpected since IQ should be expected to relate most closely to academically-related learning. However, the results suggest a very provocative hypothesis. That is, when even academic learning is optimized by formulating concrete objectives, and instructional arrangements specifically formulated to attain them, the child's response may be better predicted by the learning potential procedure, which also estimates a child's ability by offering carefully structured opportunities in which to learn prior to demonstrating one's skill on the reasoning problems. By contrast an IQ test measures what the child has learned and the child's ability to produce it upon demand in areas relevant to school success. The approach of the objectives-based curriculum and LP assessment are parallel. They both examine learning in the context of an optimized situation, presuming no necessary prior experience with the particular task or competency (in the new jargon), and present the necessary skills, with time for practice before passing judgment on the child's capability. One of the most desirable features of the objectives-based program is that the onus for training is placed in the trainer (teachers). Again, this is parallel to the manner in which learning potential assessment is carried out with children when the onus for ensuring the child understands the problem is placed on the trainer. Evidence in support of this finding that LP assessment in contrast to IQ predicts response to educational treatments has been accumulating from various sources. Budoff, Meskin, and Harrison (1971) demonstrated it using a nonverbal manipulative science unit. Budoff, Corman, and Gimon (1976) reported a similar finding with Spanish-speaking, Puerto Rican students. Budoff and Gottlieb (1976) showed differences in response to segregated and integrated placements among EMRs, although the differences in response to the mainstreamed opportunity one year after placement largely reflected motivational changes, and only small academic gains. Babad and Bashi (1975) reported that posttraining LP scores predicted attainment after teaching a cryptarithmetic unit to Jewish middle class and culturally different seventh graders in Israel. Bashi (1976), who replicated this teaching program with Israeli Arab students showed similar findings. Ployer and Gottlieb (1973) showed that posttraining Raven scores correlated with postteaching scores on a Geoboards Mathematics unit.

There has been no multi-year effort to test the hypothesis that LP scores will predict reading and mathematics attainments among special class EMRs. This extended period is required, in my estimation, because the IQ-defined EMR who has been administratively designated as such by placement in a segregated class requires one and perhaps two years to begin to cope effectively with academic materials, and be able to consolidate this learning so that it becomes evident in his everyday work, and/or his test scores.

For the child who has experienced serious difficulty in school, LP assessment provides a teacher with a basis for a more positive set of expectations for a child whose school work defines him as an inadequate student with a low scholastic aptitude (IQ). These positive expectations in turn should set in motion a drastic reorientation of the child's school program, since the child can now be seen as able to profit from learning experiences. The challenge for the teacher then becomes one of finding ways by which the child can become productively involved in school learning. Like the emphasis which is finding expression in work with perceptually handicapped children, the modes and materials of instruction can then become tuned more closely to the assets and difficulties displayed by the educationally retarded (high LP) child who has been placed in the special class for the EMR.

Some comments are required regarding the problem of understanding the learning potential differences within these IQ-defined EMR populations. It is too early to conclude that the nongainer child is mentally retarded because he has failed to profit from the training on our criterion tasks. There is certainly a group tendency for nongainers to more clearly fulfill the stereotype ascribed to mentally retarded children. Their failure to profit from training or to learn very slowly or only on low level tasks (e.g., low cognitive load versions of Jensen's level one tasks) is certainly one characteristic. Their tendency to aspire unrealistically in the face of failure, to define goals for themselves vocationally or socially they cannot realize, to see themselves as more maladjusted, and to be economically dependent as adults also tends to operationally define them through these behaviors, as mentally retarded However, there are probably misclassifications within the nongainer class.

It is quite clear that the training procedure adopted for Kohs Block Designs was arbitrarily chosen. We taught one systematic strategy for understanding the principles relevant to solution of the problems. It is quite likely that some subjects did not respond to our type of training and would respond to a different type of training. For example, Frank (1970) designed a subject-controlled training procedure for the nongainer which contrasts with the experimenter-controlled training procedure we have been using. In this subject-controlled training procedure, the S is shown 2, 4, or 9 blocks and told the restrictions in making the design--i.e., with 2 blocks, one side of each of the two blocks must touch; or with 4 or 9 blocks, that the blocks must all be placed within a 2" x 2" or 3" x 3" square. He is then asked to make as many designs as he can using the blocks. The child or the E draws each design on a grid paper so that the designs that S makes can be remembered. The latent intent of this procedure is to have the child see the transition from the three dimensional space of the blocks to the two dimensional space of a design. When the child has successfully completed a number of patterns in each category, he is asked to reconstruct his designs from the grid paper.

What is interesting is that some nongainers, as defined by the E-controlled training procedure, suddenly seem to understand what the solution of block design problems requires after having gone through this training procedure. On a post-training trial, they perform far more satisfactorily, often solving difficult nine or even sixteen block designs when they had barely solved simple four block problems after E-controlled training. At this point we have no idea of the predictive value of this improved performance and whether these children are in fact more competent than their nongainer status indicated. But it is likely that some of the misclassifications in the nongainer category may be reduced by providing training procedures more suited to the child's individual style of learning.

What is clear at this point is that the nongainer may or may not be truly retarded, nor may this be the significant question to concern us at this time. While misclassification, when it results in educational or psychological harm to the child, is a critical consideration, greater energies must be expended in seeking translations from test results to appropriate individualized educational plans. Evidence of ability to profit from a subject-paced procedure may suggest a variety of procedures by which this child can be engaged in productive learning activity that may differ from those offered children who can respond to teacher-initiated activities with a minimal verbal demand.

Finally, the test-train-retest paradigm may have broader implications for assessment and more particularly for the types of expectations these data help establish or reinforce. The hidden curriculum of the schools assumes a variety of information handling strategies as prerequisite to school success. Children are expected to manage ever greater loads of factual information, show increasing facility in memory and abstracting or mediational strategies, to cite but a few. School curricula tend, however, not to teach strategies by which one can learn more successfully, but rather factual contents. Presumably successful students acquire the appropriate skills and knowhow that facilitate educational progress spontaneously.

It is probable that the poor urban student has not spontaneously acquired the psychological processes relevant to educability by the time he enters school or as he proceeds through school. If a taxonomy of necessary or appropriate skills that facilitate educability were available, testing of the high risk child could then be focused on a functional analysis of the child's strengths and weaknesses. Training programs could be designed that would teach appropriate skills, e.g., mediational strategies for learning, "chunking" strategies to facilitate recall,etc. Post-training examinations could focus on discerning the extent to which the strategy was employed, i.e. the child's trainability, and the degree to which the newly acquired strategies are used with non-trained novel materials. Ultimately, the challenge is to have the facilitating strategy become incorporated into the child's repertoire of use-strategies in problem solving.

Potentially, then, a training-based assessment paradigm may lead to a functional diagnosis of strengths and weaknesses among high risk students. Assessment would be concerned with testing the extent to which the child can profit from appropriate training programs, for example, as in information processing strategies. The psychologist would relate his efforts to understanding the child in terms of the child's actual learning processes. He/she would seek to determine the extent to which a problem is intractable by employing training programs which are designed to aid with the particular hangup. In turn, he would continually relate this educational psychological understanding to the child's teacher and seek to translate very specifically into classroom practices, the way in which the teacher could facilitate the child's learning.

This change in emphasis would constitute the major break with our traditional concerns in testing. It would require us to de-emphasize long term predictions based on how well the child has run a particular experience track. That is, the logic underlying the Binet is based on the very powerful assumption that the child who has mastered certain facts and cognitive use-strategies spontaneously and can produce them upon demand, will continue to do so. Hence, the prediction of school success. Since so many white collar occupations depend on a modicum of success in school, the Binet-based prediction is very powerful. Because the implicit asusmptions of equal access and use of their use-strategies for requiring, storing, and producing information are violated with children from economically and culturally different backgrounds the Binet type of judgment is justly attacked as discriminatory. Since teachers' expectations are based on similar premises to those underlying the Binet, the expectations of failure for so many of these children enhances this prejudice.

Adoption of the learning potential type of assessment paradigm would alter the basic questions being asked by testers. Instead of being primarily concerned with sifting children into educational tracks, and laying the groundwork for the self-fulfilling prophecy, testing would be concerned with the child's present status in some finite area(s) of functioning, and his amenability to training on that function(s). The psychologist's contribution would be more closely related to the educational situation of the child, and might involve quite specific recommendations regarding how he might best be taught, whether the teaching is done by the classroom teacher or remedial personnel. The information relayed to the teacher is that Johnny has trouble handling ever increasing information loads and hence tends to get lost in a classroom when too many facts are being transmitted by the teacher. The teacher could then seek to develop a learning context most appropriate to Johnny by minimizing loads and, concurrently with the psychologist, help Johnny manage heavier loads by appropriate training regimens. Johnny's continuing response to training in the deficient skills would behaviorally define his ability to profit from these experiences and tend to structure expectations around Johnny's performance in a behavioral microsituation with the psychologist. More positive expectations would result since most children would demonstrate some progress. The psychologist's role then would be to develop and look for ways by which the child's difficulties could be better understood in the context of what the school's job is, i.e., educating children.

The child's positive response to training would gain the child an ally in the psychologist since the focus would be on short term alterations through training, and close working relationships with teachers so the child's shortcomings can be translated into relevant classroom procedures. The focus on helping the child attain a desired educational status in this stepwise manner would tend to minimize the negative effects of the long range determining statement, i.e., the IQ, that results in the self-fulfilling prophecy. The child who tends to respond minimally to the training opportunities and functions poorly in many other nonacademically dependent situations would define himself as mentally retarded, since behaviorally he manifests a marked slowness in profitting from experience.

In summary, then, what we require to break the cycle of the self-fulfilling prophecy is a school system oriented toward optimizing education for each child. For poor and/or culturally different children, it requires a de-emphasis on the predictive statements represented by IQ scores which tend to sift children into implicit or explicit tracks. It requires a greater multiplicity of structural-administrative options by which educationally needy children can be taught than usually exists in urban systems today. It demands a broader range of equally acceptable paths by which children may attain an educated status, and a committment, philosophically, by school administrations to insure successful educationally meaningful achievement for every child.

Our data using the learning potential strategy suggests also that assessment procedures for high risk children should be restructured so that we stop reifying his failure to come from a middle class home into a long term prediction of school failure. Rather, structuring assessment around short-term empirically verifiable questions by using an amenability to training model would diminish the need for the long term prediction. More importantly, this type of testing model would provide teachers and remedial personnel with information that could be made educationally relevant to their day to day work with the children. The ultimate assessment of children's capability would be made by the child's response to training, rather than on the kinds of experiences and skills which children should have spontaneously acquired prior to entry into public school or early in their school career. If the poor urban child from a disadvantaged minority group background seems to have skills that are not congruent with those required by the school's middle class bias then it is incumbent on the schools to recognize this discrepency and seek ways by which to close the gap.

3rd

The assessment process suggested, in its sketchiest outline, demands serious attention by psychologists since it demands an analysis of the psychological processes that facilitate educability. If we can identify the critical processes, the process of understanding how to train these processes is probably very simple. The most difficult aspect of this training process is to learn how the deficient process can be integrated into the child's repertoire of use strategies.

Finally, the most critical change this approach dictates is that the child's failure to learn must be perceived as the responsibility of the school and its professional personnel since they are charged with educating children. The most lamentable feature of many publicly supported institutions is the ease with which they assign responsibility for failure to their clients' failings rather than to their own lacks. If our educational and mental health institutions would take more active responsibility for their clients' failures, we would have taken the largest step toward true accountability.

References

Babad, E. & Bashi, J. Final report: An educational test of the validity of learning potential measurement. Ford Foundation, June, 1975. *RIEP-Print* No. 91.

Babad, E. & Budoff, M. Sensitivity and validity of learning potential measurement in three levels of ability. *Journal of Educational Psychology*, 1974, *66*, 439-447.

Bashi, J. & Babad, E. Learning potential assessment with Israeli Arab students. *RIEP-Print* No. 92, in preparation.

Baumeister, A. A. & Kellas, G. Intrasubject response. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 3). New York: Academic Press, 1968.

Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. *Journal of Personality*, 1961, *29*, 303-320.

Budoff, M. Learning potential among the educable mentally retarded. Progress report to the National Institute of Mental Health, 1965 (mimeographed).

Budoff, M. Learning potnntial among institutionalized young adult retardates. *American Journal of Mental Deficiency*, 1967, *72*, 404-411

Budoff, M. Learning potential as a supplementary strategy to psychometric diagnosis. In J. Hellmuth (Ed.), *Learning disorders* (Vol. 3). Seattle, Wash.: Special Child Publications, 1968.

Budoff, M. Learning potential: A supplementary procedure for assessing the ability to reason. *Seminars in Psychiatry*, 1969, *1*, 278-290.

Budoff, M. Learning potential and institutional discharge status among young adult EMRs. *RIEP-Print* No. 33, 1972.

Budoff, M. Learning potential and educability among the educable mentally retarded. Progress report on Project No. 312312 submitted to the National Institute of Education, Feburary, 1973. RIEP-Print No. 72.

Budoff, M. & Corman, L. Demographic and psychometric factors related to improved performance on the Kohs learning potential procedure. *American Journal of Mental Deficiency*, 1974, *78*, 578-585.

Budoff, M., Corman, L., & Gimon, A. An educational test of learning potential assessment with Spanish-speaking youth. Interamerican Journal of Psychology, 1976, 10.

Budoff, M., & Friedman, M. Learning potential as an assessment approach to the adolescent mentally retarded. Journal of Consulting Psychology, 1964, 28, 434-439.

Budoff, M., Gimon, A., & Corman, L. Learning potential measurement with Spanish-speaking youth as an alternative to IQ tests: A first report. Interamerican Journal of Psychology, 1974, 8, 233-246.

Budoff, M., & Gottlieb, J. Special class EMR children mainstreamed: A study of an aptitude (learning potential) X treatment interaction. American Journal of Mental Deficiency, 1976, July.

Budoff, M., & Litzinger, S. A sample set and brief manual for learning potential assessment using Raven's Progressive Matrices. RIEP-Print No. 59, 1975.

Budoff, M., Meskin, J., & Harrison, R. H. Educational test of the learning potential hypothesis. American Journal of Mental Deficiency, 1971, 76, 159-169.

Budoff, M., & Pagell. W. Learning potential and rigidity in the adolescent mentally retarded. Journal of Abnormal Psychology, 1968, 73, 479-486.

Budoff, M., & Pines, A. Reaction to frustration as a function of learning potential status. RIEP-Print No. 9, 1970.

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measures. RIEP-Print No. 56, 1973.

Corman, L., & Budoff, M. A comparison of group and individual training procedures on the Raven learning potential measures with black and white special class students. RIEP-Print No. 57, 1973.

Corman, L., & Budoff, M. Factorial composition of the Raven Progressive Matrices in the context of learning potential assessment. RIEP-Print No. 62, 1973/1975.

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven Progressive Matrices. Educational and Psychological Measurement, 1974, 34, 407-412.

Corman, L., & Budoff, M. Demographic and psychometric factors related to improved performance on the Raven learning potential procedure. RIEP-Print No. 42, 1975.

Corman, L., & Budoff. M. IQ and learning potential measurements of general intelligence: A comparison of relationships. In D. A. A. Primrose (Ed.), Proceedings of the Third Congress of the International Association for the Scientific Study of Mental Deficiency (Vol. 1). Stirlingshire, Scotland: IASSMD, 1975.

Elementary Science Study of Education Development Center. Batteries and Bulbs (Trial Teaching Edition). Newton,MA: McGraw-Hill, 1966.

Feuerstein, R. Learning potential assessment device. In B. W. Richards (Ed.), Proceedings of the First Congress of the International Association for the Scientific Study of Mental Deficiency in Montpelier France. Surrey, England: Michael Jackson Publishing Co., 1968.

Folman, R., & Budoff, M. Learning potential and vocational aspirations of retarded adolescents. Exceptional Children, 1971, 38, 121-130.

Folman, R., & Budoff, M. Attitudes toward school of special and regular class adolescents. RIEP-Print No. 32, 1972.

Folman, R., & Budoff, M. Learning potential and family status among special (EMR) and regular class adolescents. RIEP-Print No. 37, 1972.

Folman, R., & Budoff, M. Social interests and activities of special and regular class adolescents as compared by learning potential status. RIEP-Print No. 36, 1972.

Frank, L. Clustering performance and recall efficiency. A normative study and some comparative data from educable retardates. Working paper, 1970, Research Institute for Educational Problems.

Frank, L., & Riley, L. A further analysis of paired associates learning and learning potential status. RIEP-Print No. 5, 1969.

Glaser, D. An analysis of paired associates data by learning potential status. Unpublished report, Research Institute for Educational Problems, 1966.

Goldstein, K., & Sheerer, M. Abstract and concrete behavior. Psychological Monographs, 1941, 53 (2, Whole No. 239).

Guthrie, G., Butler, A., & Gorlow, L. Patterns of self attitudes of retardates. American Journal of Mental Deficiency, 1961, 66, 222-229.

Hamilton, J. L. Application of the learning potential paradigm to severely mentally retarded adolescents. Unpublished doctoral dissertation, University of Missouri, 1972.

Hamilton, J. L., & Budoff, M. Learning potential among the moderately and severely retarded. *Mental Retardation*, 1974, *12*, 33-36.

Harrison, R. H., & Budoff, M. A factor analysis of the Laurelton Self-Concept Scale. *American Journal of Mental Deficiency*, 1972, *76*, 446-459.

Harrison, R. H., & Budoff, M. Demographic, historical, and ability correlates of the Laurelton Self-Concept Scale in an EMR sample. *American Journal of Mental Deficiency*, 1972, *76*, 460-480.

Harrison, R. H., Greenberg, G., & Budoff, M. Differences between educable mental retardates and nonretardates in fluency and quality of verbal associations. *American Journal of Mental Deficiency*, 1975, *79*, 383-591.

Harrison, R. H., Singer, J., Budoff, M., & Folman, R. Level of aspiration as a function of learning potential status in the educable mentally retarded. *Psychological Reports*, 1972, *30*, 47-57.

Hurtig, M. Experimental study of the possibility of intellectual learning in normal and retarded children. *Enfance*, 1959, 371-383.

Hutten, L. A test of verbal comprehension in retarded children. Report, Research Institute for Educational Problems, 1973.

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills by first graders. *Child Development*, 1971, *42*, 149-159.

Jensen, A. R. Scoring the Stroop test. *Acta Psychologica*, 1965, *24*, 389-508.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*. 1969, 39, 1-123.

Kagan, J. Individual differences in the resolution of response uncertainty. *Journal of Personal and Social Psychology*, 1965, *2*, 154-160.

Kohs, S. C. *Intelligence measurement*. New York: MacMillan, 1923.

Kounin, J. Experimental studies of rigidity: I. The measurement of rigidity in normal and feeble-minded persons. Character and Personality, 1-41, 9, 251-273.

Lewin, K. A dynamic theory of personality. New York: McGraw-Hill, 1936.

Mandelkorn, B. The development of a communication training technique for young EMRs from age seven to eleven. Unpublished doctoral dissertation, Harvard Graduate School of Education, 1973.

Mandelkorn, B., & Corman, L. The effects of two training models on communication-related abilities and intelligence in young, IQ-defined EMRs. Paper presented at the Society for Research in Child Development meeting, Philadelphia, March, 1973. RIEP-Print No. 45.

Mankinen, R. Discrimination learning of EMRs as a function of learning potential and striving orientation. RIEP-Print No. 17, 1970.

Mankinen, R. The relationship of an impulsivity measure to learning potential status. RIEP-Print No. 12, 1970.

Mankinen, R. Learning potential status and verbal mediation training with EMR students. RIEP-Print No. 29, 1972.

McCaffrey, A. The minimal distance principle in syntax reconsidered: Development of comprehension of a complex grammatical construction among normal and educable mentally retarded children 7 to 12 years of age. Paper presented at the Society for Research in Child Develop= ment meeting, Philadelphia, March, 1973. RIEP-Print No. 61.

Mercer, J. Sociological perspectives on mild mental retardation. Paper presented at the First Annual Study Conference in School Psychology, 1972.

Pines, A., & Budoff, M. Studies in personality correlates of learning potential status in adolescent EMRs: reaction to frustration and self-peer evaluation. RIEP-print No. 16, 1970.

Ployer, F., & Gottlieb, J. A pilot study to investigate the use of Geoboards to validate the group learning potential assessment with first grade children. RIEP-Print No. 46, 1973.

Raven, J. C. Coloured progressive matrices, sets A, Ab, B. London: H. K. Lewis, 1956.

Raven, J. C. *Standard progressive matrices, sets A, B, C, D, E.* London: H. K. Lewis, 1958.

Sarason, S. B., & Gladwyn, R. Psychological and cultural problems in mental subnormality: A review of research. *Genetic Psychology Monographs*, 1953, *57*, 3-289.

Suzman, S. A comparative study of normal and EMR speech performance on a paraphrase task. Unpublished report, Research Institute for Educational Problems, in preparation.

Turner, R. H. *Social context of ambition: A study of high school seniors in Los Angeles*. San Francisco: Chandler Publications, 1964.

White, B. An initial validation of the Raven's learning potential procedure. Unpublished report, Research Institute for Educational Problems, 1968.

Zigler, E. Social deprivation and rigidity in the performance of feeble-minded children. *Journal of Abnormal and Social Psychology*, 1961, *62*, 413-421.

Zigler, E. Rigidity in the feeble minded. In E. Trapp & P. Himmelstein (Eds.), *Readings on the exceptional child*. New York: Appleton-Century-Crofts, 1962.

APPENDICES

Appendix A	Raven Learning Potential Test: Manual Directions for Administration Scoring Key and Norms Answer Sheets
Appendix B	Orientation Training for Raven Progressive Matrices: Instructions Answer Sheet Test Booklet
Appendix C	Raven Learning Potential Test: Directions for Administration--Spanish Version
Appendix D	Raven Learning Potential Test--Adapted for Moderately and Severely Retarded
Appendix E	Picture Word Game: Instructions Training Booklet Test Booklet

APPENDIX A

Manual for the Raven Learning Potential Test

Part I. Overview of the Learning Potential Procedure

The Learning Potential (LP) assessment comprises a test-train-retest procedure. A substantial fraction of poor school learners show a marked improvement in competence on reasoning tasks when they have had opportunities to work with the test materials within the context of problem-relevent training. This manual presents training procedures oriented toward increasing the child's understanding of problems which appear on the Raven Progressive Matrices (RPM). The Learning Potential assessment oriented around the RPM requires that this test be administered twice, prior to and following the training. The standard test administration procedure is drawn from Raven's manuals and is presented below. Included also are the instructions for providing problem-relevant training in small groups for students ranging in age from 7 years to adulthood. A separate version exists which is designed for children between 5 to 7 years of age. With suitable alterations of wording, the procedures may also be employed with individual persons. It should be noted that LP assessment procedures are also available for moderately and severely retarded persons employing the Raven Progressive Matrices in a formboard procedure, and with the Leiter International Scale (See Hamilton & Budoff, 1974; Budoff & Hamilton, 1976). Studies with Puerto Rican and Mexican-American students of average IQ or less have shown a learning potential approach to be appropriate when used with Spanish-speaking slow learners and have demonstrated the predictive validity of LP assessment for these non-English-speaking students. (See Budoff, Gimon & Corman, 1974; Budoff, Corman & Gimon, in press.) These instructions for testing and training are available in a Spanish translation.

The training procedures which are generally completed in two half-hour sessions on two successive days are oriented toward helping the child(ren) understand what they must do in order to solve the various types of problems appearing in the RPM. One of the difficulties in developing the instructions for the training is that the manner in which different children come to understand the same concept differs. In previous versions of these training instructions, we have sought to emphasize the teacher's obligation to ensure that the child comes to understand the relevant principle(s) contained within a problem if he can. We stressed multiple administrations of the materials in somewhat altered presentations so that this might be possible. Those instructions were written for individual administration of the training. Due to the requirements of the group administration to several children concurrently, the options open to the teacher have had to be narrowed. The wide variety of options for presentation of the principles and concepts applicable to a particular problem are not indicated in this version because of the strictures of group administration.

In these instructions, we have strived to emphasize a strategy of training that encourages active participation by the students, minimizes verbal explanation, and maximizes visual aids presented within an active teaching role, e.g., presenting the stimuli and the choices on a blackboard via 35 mm slides so that the child may indicate his(her) choice to the teacher who draws it in. We suggest the teacher (T) choose children to respond who may produce incorrect responses so that these examples may serve to highlight for the children the relevant attributes of each problem the children should attend to. But these choices must be handled within a non-punitive context. That is, the choice offered by a child should not be publicly called wrong even if incorrect. It should be taken as a good faith offering of a child who wants to cooperate in a group learning situation. The adequacy of the choice should be evaluated by having T draw it on the blackboard, in the missing space provided, asking the child who provides the answer whether this is how he/she saw it. Collectively with the other children, the teacher should encourage the other students to decide whether the choice is appropriate. If the children do not feel the choice is adequate, they should be asked in what ways it falls short of a correct choice for that pattern. These situations serve as "live" examples by which to illustrate the critical attributes of the problem the children should consider. The more instances of this kind the teacher can generate from the children themselves, the livelier the training session will be. The training procedure also includes specific instructions on some problems to illustrate the inappropriate features of certain choices for particular problems.

The major worry about this procedure is that the teacher not slow down the training procedure unduly such that the slow pace becomes tedious to the participants. One cannot emphasize too strongly that the teacher must maintain a sprightly pace during the training procedure so that the attention of the group will not flag. Hence, the emphasis on not explaining too much and involving the students in the process of figuring out the problems. We have attempted to write the instructions for the training using short, declarative sentences, and language which involves the students in the process of solution. PLEASE PAY MOST ATTENTION TO MAINTAINING A STEADY PACE THAT IS NOT VERBOSE AND EXPLANATORY! FAILURE TO DO SO REPRESENTS THE FASTEST WAY TO LOSE THE ATTENTION OF THE GROUP, AND RESULTS IN

BEHAVIOR PROBLEMS. ADULTS TEND TO TALK TOO MUCH!! It is also extremely important that the teacher attend to the child(ren)'s level of fatigue as the critical cue for continuing the training. Signs of fatigue may include eye rubbing and signs of sleepiness, but more likely, a rising level of irritability that becomes more and more difficult to control.

To repeat, the intent of this procedure is to involve the children in the learning process, and to use the products of the children's efforts to illustrate and highlight the particular features the child should attend to in solving these various types of problems--whether these are orientation of the elements to each other in a particular pattern, or the completion of a design throughout the figure. It should not be used to chastise or otherwise belittle the child's well-intended contribution to the group learning process.

The primary intent of the training is to demonstrate in many different ways that the answers to the various types of problems depend on the student's active search for a principle and that these principles are not arbitrary or capricious. The teacher must always communicate that there is a definite pattern or principle(s) by which one can find the correct solution that is dictated by the dimensions set forth in the problem. Solutions are derived by analyzing the problem elements; deciding how they relate to each other. Once this is done, the solution must follow. Hence, the principle(s), if it (or they) are correct for the problem, must be shown to be applicable within the problem in more than one way. That is, a correct principle can be checked by using it to solve the problem in another way. As you will see, this particularly applies in the classification problems in which a solution which can be figured out by working on the problem one row at a time should also work when the problem is solved using one column at a time.

Part II. Testing/Training

Objectives

The training procedure is designed to teach the children to attend to all the choices presented in each problem and to think about the problem before responding rather than impulsively choosing an answer. This training procedure should be adhered to as closely as possible. In some instances, if the teacher encounters difficulties with some children's understanding of the problem, he/she may use other examples which may be helpful to the child's understanding. But the teacher must balance the need to help all children understand the particular problems against the problems of proceeding so slowly that the children who have grasped the problems more quickly may become restless and create disturbances. To keep the group working and learning as a unit, it is critical to recognize that the explanations we feel compelled to make because we wish to make sure the child understands all the possible sources of errors he/she can make, may not be seen as relevant by the child.

The training procedure helps children deal more effectively with four types of problems found in RPM Sets A, AB, B: discrete pattern completions, continuity and reconstruction of simple and complex structures, reasoning by analogy in a 2 X 2 format, and easy linoleum block completions. The procedure allows alternative approaches to be presented to the student to make sure that he understands the requirements of the problem.

What we had found was that children who did solve the 2 X 2 problems with ease by the close of the training period did not spontaneously see that a similar solution process could be applied to the 3 X 3 format found in RPM Sets C, D, and E, which had not appeared in the context of the training. Thus, there was very little transfer of the training on the 2 X 2 to the 3 X 3 problems, especially among the slow learning population. Training on three principles appropriate to 3 X 3 problems attempts to direct the child's attention to the fact that these problems are similar to those already mastered and can be solved with some focused effort.

The 3 X 3 training designs are included to help the student see the pertinence of the training on the double classification problems to another type of problem, namely 3 X 3 problems. Also, since three principles of solution are included in the 3 X 3 set of training designs, the child comes to understand that there are different principles that can be applied in solving these 3 X 3 types of problems and comes to see that he must identify the principle(s) that work(s) for the particular problem. He also learns to check whether his initial hypothesis aboutthe appropriate principle(s) work(s) for the proble

The problems are presented to help the person see that what he learned in solving the double classification problems is pertinent to a different problem format. He comes to see that he must choose a principle, test its applicability, and then check whether it works by solving the same problem in an alternate manner.

Throughout the training period, a number of principles are adhered to:

1. No problem which appears in Raven's Progressive Matrices is presented in the training series so that there is no item-specific training.

2. Great efforts are made to be certain that the children understand the demands being made of them on the task.

3. The requirements of the task are demonstrated initially on materials that are pictorial and meaningful, i.e., an American flag and flowers for the pattern completion tasks. Successive problems require the children to deal with geometric forms, and so adhere to the basic format of the Matrices Test stimuli.

4. In devising training for double classification problems it was found that children could easily derive one attribute at a time but often did not hold the first attribute in mind while they derived the second relevant attribute. During development of this training procedure, the child's understanding was facilitated by drawing the relevant attributes, one at a time in a training booklet and on the blackboard, as he derived them. This helped concretize the elements of the solution process so that many children, after this type of practice, could do the double classification problems without "note-taking," in their heads.

5. The training procedure is wordy and can be difficult when working with low school achievers or mildly mentally retarded children. The teacher should encourage class participation, accept any choice as a possibly correct one, and illustrate by chalk drawing (rather than explanation) whether or not the choice would complete the design. Training should be action-oriented, seeking to involve the students as much as possible.

Method

Ideally, for group training, the training session should be scheduled the day after the pretest unless children are grouped by pretest score. The posttest should be the day after training is completed. The training procedure described is for small groups, e.g., three to fifteen children depending on age and ability level. It can also be used for individual children with suitable minor adaptations of the instructions. We have found the optimal group size for children with IQs in the borderline or mildly mentally retarded ranges to be no larger than six.

Students with higher IQs can be seen in larger groups for testing and training, although except in unusual circumstances these groups should not exceed 15 members. In this latter situation, it is desirable to have an observer who can circulate among the students to be certain that they are working on the problems successfully. This person can offer individual help to a student who is floundering, or who has lost his/her place in the testing or training booklet without disruption to the flow of work for the other children. A major factor in determining the size of the group is the capability of the teacher/tester to manage the testing or training process such that the students work without disruptions.

A useful procedure to maximize the work with groups of slow learning students is to group the students for the training groups by their pretest Raven score. Be certain that all the students have been administered the same sets of problems in calculating the scores and in making group assignments for training. The procedure would involve administering the pretest with larger groups (with an aide to make sure the students record their answers properly, i.e., keep their place on the answer sheet. These answer sheets should be scored, and students assigned to groups relatively homogeneous in pretest level. The assumption is that students with the same approximate level of achievement are experiencing difficulty with comparable levels of difficulty of the problems. While this assumption can not be shown to be completely accurate, grouping by pretest competence level should simplify training by reducing the heterogeneity among the group members. The major deviation from the recommended procedure when using this grouping technique is an increase in the number of days between pretest and the start of training. This interval is the least critical aspect of the recommended procedure. The interval between the training sessions and between training and posttest should be more closely controlled and should be carried out on successive days.

Children who are unruly and cannot be controlled or cannot keep up in the group administration can be handled in two ways. If there is an aide present, the aide can work with the child individually, occasionally moving away from the child to check the progress of the other children in the group. Close attention to the child may make it possible for the child to maintain control over his behavior and/or to progress more satisfactorily. If no aide is present or available, the child should be separated from the group, returned to his classroom, and seen subsequently for an individual or smaller group training experience. One should take care, however, to be certain that the particular child be separated because he/she can seriously disrupt the work of the other children.

When it is suspected that students may have difficulty coordinating the test answer sheet and the test booklet when recording answers, sample set S should be administered. First graders, young EMR children, and learning disabled children are potential candidates for needing practice to fill in the answer sheet properly. The teacher should consult with the classroom teacher as to whether this set should be administered since accurate test results demand that the child place his answer in the appropriate space on the answer sheet. Ideally, the sample set is administered by the classroom teacher or the trainer the day before the Raven pretest although the trainer can administer Set S immediately before the pretest. Ten minutes are usually adequate for administering this 6-page booklet. Each page shows pictures of six objects in the same format as the 6 pieces below each puzzle in the Raven A, AB, B booklet. Each piece is numbered. The child is instructed to find one of the 6 objects on each page, then transfer its number to the proper place on the enlarged spaced of the sample answer sheet. Appendix A contains the instructions and the stimulus figures.

Materials and Procedure

Equipment

The instructor will need a carousel projector with remote control, a blackboard and chalk. A blackboard is a good surface for projecting the pictures. It is used by the teacher or a child to draw in the missing parts of the designs, to make the correct choice for a problem, and allows the children to see whether or not their choice was correct. In this way, the teacher can provide graphic illustrations as part of the training.

Materials

Pencils and erasers should be brought by the children to the test sessions, although the instructor should have extras for emergencies. The training procedure uses slides projected on a blackboard surface. A carousel slide projector is generally used.

Procedure

Generally, there are two slides for each puzzle: the first one follows the same format as the Raven puzzles, is exactly like the picture in the child's book; the slide following this stimulus slide has the correct answer filled in. However, not all training designs appear in the child's booklet. The several designs which do not appear are coded with asterisks (**) in the training instructions. Identification information in the training instructions follows the following format:

1. asterisks (**)(if applicable)
2. design label
3. slide number(s)
4. page number (if applicable)

In order to avoid unnecessary confusion the teacher must be thoroughly familiar with the training materials and the instructions for presenting them. The usual sequence in presenting each item in training is as follows:

1. show slide
2. call on a child for choice
3. draw in piece indicated by child so he can actually see on the blackboard what he visualizes in his mind or has drawn in the booklet.
4. ask group if they agree with choice given
5. solicit other choices from children who don't agree. Draw in the option(s) given on the blackboard and solicit opinions from the rest of the group as to the adequacy of the choice. The point of this procedure is to actively involve the children in solving the problems so as to force them to attend to the critical details by examining them within the context of solving actual problems.
6. show answer slide as a final check to see if what all agreed on is the correct solution.
7. change slide to next design.

Encourage children to keep with you during training--stress looking up at the blackboard--booklets should only be used for solving certain problems and recording responses which may be derived by group activity. This should assure that no children jump ahead, and that all end up with the correct response.

It is especially important during the first few 2 X 2 problems that the teacher's explanation is clear and concise and that the group is attending to the instruction. It is critical for the children's understanding of the double classification problems that the teacher has a clear understanding of the rationale for explaining the solution. The training instructions for each problem follow the same basic format. The two attributes of the missing piece (shape and size) are solved for separately. The correct answer is the merging of these attributes. Shape is always solved for first so it varies whether the child must look at the problem horizontally or vertically first. The concept of change is stressed and children are instructed to look in two directions for changes. Until the last few simple 2 X 2 problems (designs 38 and 39), the teacher always draws the shape into the empty space of the puzzle at the board so the children do not have to keep both attributes in mind at once.

Three different children may be called on to respond to each puzzle. The first child gives the shape, the second child the inside, and the third child finds the correct answer below the puzzle. Regardless of whether or not the child's idea for the shape of the missing piece is correct, the trainer draws it in, states the rule for changes, and gets the group consensus of the correctness of the choices.

Part III. The Assessment Procedure

Synopsis of Sessions

Sample Set S is administered before the pretest if necessary. If there is some question as to whether the children can easily master the intricacies of the redesigned answer sheet for the Raven Progressive Matrices, then the sample booklet of problems should be administered first, either in a separate session or as a preliminary to the initial administration of the RPM. The key consideration is the chronological age of the children, or the time they can expend in useful effort. The younger the children, or the more restricted the time for sustained output, the more likely that the conservative approach requires that the sample booklet be administered at a separate session.

Session 1. Administer age-appropriate sets of the Raven Progressive Matrices in groups. If large numbers of children, heterogeneous in ability, are being tested and trained, it would be desirable to score these tests and assign children to training groups who have relatively similar pretest scores when the same sets of problems have been administered to the children. Children who cluster at the same approximate score levels will require similar kinds of help. A heterogeneous group in which children who understood relatively harder concepts are grouped with those who understood only the simpler problems poses potentially difficult problems related to wavering attention, which can lead to a decay in the teacher's hold on the group. Discipline and attentional problems may arise more frequently in the heterogeneous groups since some children will wish to go much faster than the slowest child in the group will permit. Control of the time interval is least critical between the pretest and the later stages of the assessment procedure.

Sessions 2 and 3. The two training sessions should be scheduled on successive school days, preferably without an intervening weekend. Our data suggest that the training is maximally effective with mildly mentally retarded children when the group size is no larger than six children. It usually requires about 30 minutes for each training session.

Session 4. Posttest on the same subsets on the Raven Progressive Matrices as administered during the pretraining phases should occur within two days of the completion of training.

Pre- and posttesting administrations use redesigned Raven answer sheets in which answers to Sets A, AB, B, and C and D are recorded on separate sheets. The answer spaces for the three scales are larger on the redesigned answer sheet and in larger type so that the children will be less confused.

References

Budoff, M., Corman, L., & Gimon, A. An educational test of learning potential assessment with Spanish-speaking youth. *Interamerican Journal of Psychology*, in press.

Budoff, M., Gimon, A., & Corman, L. Learning potential measurement with Spanish-speaking youth as an alternative to IQ tests: A first report. *Interamerican Journal of Psychology*, 1974, *8*, 3-4.

Hamilton, J. L., & Budoff, M. Learning potential among the moderately and severely mentally retarded. *Mental Retardation*, 1974, *12* (4), 33-36.

DIRECTIONS FOR ADMINISTRATION OF RAVEN MATRICES

After pupils are seated distribute pencils and answer sheets. The Instructor says, "On the table in front of you, you will find an answer sheet. Please write your name, both first and last names, at the top of the page." Red booklets (Sets A, A_B, B) are then distributed and pupils told not to open the books until everyone is ready.

The instructor then says, "Open your books to the first page with the green puzzle. It is like this (shows his book). At the top of the page it says Set A and you have a column A here on your answer sheet (points to sample on blackboard). This is problem A.1. You see what it is. The upper part is a puzzle with a piece missing. Each of these pieces below (instructor points to each in his book) is the right shape to fit the space, but they do not complete the pattern. Number 1 (T points to the piece and then to the pattern) is the right shape but the wrong pattern. Numbers 2 and 3 are wrong--they fit the space, but they are not the right pattern. What about number 6? It is the right pattern (instructor illustrates that the pattern is the same as the pattern above) but it does not go all over (T points to the white piece). Put your finger on the one that is completely right. Yes, number 4 is the right one. So the answer to A.1 is 4. Write 4 here (points to answer sheet on blackboard), beside A1. on your scoring form. "Do not turn the page over yet." "Are there any questions?"

"On every page in your book there is a puzzle with a piece missing. You have to decide each time which of the pieces below is the right one to complete the puzzle. When you have found the right piece, write its number on your answer sheet beside the number of the pattern. They are simple at the beginning and get harder as you go on. There is no catch. If you pay attention to the way the easy ones go, you will find the later ones 'not as hard.'"

"Let's try another one together. Turn to the orange page to A2. Look at the big pattern at the top of the page to find the piece that is missing. Now, look carefully at all the six pieces below. Remember only one is totally correct. Pick the one which you think is completely right and write its number beside A.2. on your answer sheet.. . . . (When everyone has it). The right one is number 5. See that you have written the figure 5 beside A.2. on your form. Are there any questions?"

"Now you can finish column A by yourself. Make sure you do twelve (12) puzzles--don't skip any pages. When you finish Column A, go back and check your answers. When you're sure they're all correct, stop and raise your hand. Work at your own pace. You can have all the time you want. Do not leave any out. Work slowly and look carefully at all the pieces below the puzzle pattern before you choose one. See how many you can get right."

(When each child raises his hand at the end of set A, the instructor should go to the child and check to see that the answers have been recorded in the correct spaces on the answer sheet. The following instructions are then given to each group.)

"Turn to the page where it says Set A_B on the top, problem

AB 1. The answer to AB 1 goes up here in the second column. Now continue doing the puzzles in Set A_B. Do all 12 problems and make sure you put your answers in the second column next to the numbers of the puzzles. When you finish this column, raise your hand.

(Repeat above procedure for Set B).

During the administration of the test there is no time limit. When a child completes each set he is reminded that he can check his answers before he raises his hand. When the children are finished, they are instructed to insert their answer sheets inside their booklets and to close their booklets. When most of the children have completed the test, the booklets can be collected from those who have finished, and the children should quietly take out some other work. Children who are really working very slowly should be encouraged to try to finish. Usually 30 minutes is adequate time to do all the problems in the red book.

The instructor must circulate constantly to check that each student is working at the correct place on the answer sheet. If a child is not on the right item, place a mark at that point on the answer sheet and redirect the student to the line which agrees with the problem he is working. At some point, review each problem with student individually to be certain that the answers he entered on each line are those he intended. This may be done during the testing period while waiting for all the students to complete their work, or more usually, at the end of the period. This is a most important part of the procedure since student's understanding can only be gauged by the answers he records. Instructor must circulate constantly to be certain that each student is writing the answers on the correct line on his answer sheet which corresponds to the problem in the test booklet.

Students who are fourth graders or older continue with Sets C and D. Set E is only given to persons above 12 years of age. Collect red test booklets and completed answer sheets and pass out tan booklets (1956) which contain Sets C, D, and E. A useful maneuver to minimize confusion is to distribute the booklets containing Sets C, D, E with the answer sheet interleaved at problem C1. Students are asked to write their names, school, and the date on the answer sheet and then are asked to close the test booklet with the answer sheet left an at C1, and wait until everyone is ready. Students are instructed not to open the books until everyone is ready.

Since the tan books (Sets C, D, E) are only given to children in the fourth grade or older, it would be best to schedule separate testing sessions for the older and younger children if a class has mixed age groups.

Instructor: "Open the booklet to where it says Set C." (Check that all students have skipped Sets A and B and are in the right place.) "You can see that this is just like what you've been doing except that there are eight pieces below instead of six. Let's look at C1. You see if you start up here (top left and go down column), this one, is to this one, is to this one, as (move finger to second column) this one, is to this one, is to this one, as (third column) this one, is to this one, is to which one? (Point to each of eight pieces in turn.) Yes, number 8 is right. Beside the number in column C on your answer sheet, write the number 8. Now you can continue through Set C. Find the missing piece in the big pattern above. Be sure to look at all the eight pieces below and choose carefully the one that is exactly right. Write its number on the answer sheet beside the number of the pattern. Work slowly and carefully. Make sure you turn only one page at a time. Make sure you do all 12 puzzles in Column C. Notice there are puzzles on both sides of every page, so make sure you look at the backs of the pages. See how many you can get right. These are more difficult so don't get discouraged. Do the best you can. You are not expected to get all of them right. When you have finished Column C, check your answers and then raise your hand." Children should be individually instructed to turn to Set D and reminded that they do not do Set E after the instructor checks that the answers are in the correct spaces. For most older students (CA > 12), the pauses between sets may be omitted and the students instructed to complete the book. Students over 12 years of age should continue onto Set E. The red books with the Colored Progressive Matrices are always collected before the tan books containing Sets C, D, and E are distributed. The children should be cautioned to make sure the numbers for the answers are always put on the line which corresponds with the number of the puzzle they are working on.

Encourage students to do their best and indicate that they may skip problems they don't understand. They should be encouraged to guess at the correct answer as there is no penalty for guessing.

For re-administration following training: Instructor: "You may remember these puzzles. You worked with them before and today you are going to have another change to show how well you can figure them out." Remember to use the tricks we learned together when you work on these problems." (Administer as indicated for the pretest.

Specific Instructions for Training Session

The projector should be set up before the children arrive. The first slide should be focused on the blackboard and the projector turned off. After the children are seated with sharpened pencils and erasers, the instructor should introduce himself/herself to the children and explain to them what they will be doing. The instructor should attempt to learn the children's names if he/she does not already know them.

(Holding up red Raven book): *Do you all remember the puzzles you did yesterday? Today, I'm going to show you some pictures like the puzzles in here and you will learn some tricks that will help you understand how to do these puzzles* (hold up red Raven's book) *better when you try them again tomorrow. Now, I will give you a booklet of your own so we can all work together to understand how to do these puzzles better. You can draw your answers in this book and when we are finished, you can take them home with you.*

(Pass out the workbooks): *Write your name on the front page of the book. This is your book to keep and I do not want you to lose it. Don't open the books yet. I will tell you when it's time to turn the page.* (Now turn to the first slide.)

What we are going to do is learn some strategies or tricks to use in better understanding the puzzles you worked on yesterday. You did pretty well and now I am going to show you how you can do them even better the next time.

DESIGN 1. (Flag, Slides #1 and 2, page 1) Here's a flag with a piece missing. To finish it, we need part of this black stripe (chalk it in), this white stripe and this black stripe. (Draw chalk arrows to indicate where stripes are missing as they are mentioned and draw in the corner). Look down here (point at each of the six choices) and cross out the piece that finishes the flag. Who can tell me which one to cross out up here? (choose a child) (Child's name) says it's (#5). Does everyone agree? Let's see if we're right. (Change to answer slide) Yes, (#5) does finish the flag. Why is this one (point to #1) no good? (Ask children until child tells you that the stripes go up and down instead of across) Why is this one (#3) no good? Right, it has to be stripes--not stars. Remember, you must always look at all the choices when you do these puzzles. (Change slide)

**DESIGN 2. (Patch of flowers, slides #3, 4). Look at this design up here. You don't have this one in your book. Who's going to tell me which piece to cross out down here? (Put X on piece child chose) Does everyone agree? Okay, let's check our answer. (Show answer slide) Good. Why isn't it this one (#4)? That's right. There are too many flowers and it's the wrong pattern. Why not this one? (Point to #6). The flowers in the puzzle are white with black centers and these are black with white centers. (Change slide)

**Indicates that the design does not appear in the training booklet.

**DESIGN 3. (blocks of 4 flowers, slides #5, 6). Let's do this one together. This design is not in your books. Which one of the six choices is the best piece to finish this puzzle? (Put an X on piece child chooses) Let's check and see if you're right. (Change to answer slide) Good. Why is (#2) the best piece? (Elicit from children that it's the only one that completes the pattern) (Change slide)

**DESIGN 4. (One flower per block, slides #7, 8). Look at this puzzle. Which one should I cross out? (Choose a child) Does everyone agree? Okay, let's see if we're right. (Show answer slide) (Change slide)

DESIGN 5. (large flower, slides #9, 10, page 2). Turn to page 2 in your books. Look at all the six choices below and find the piece to complete this puzzle. Mark it with an X. (Circulate quickly to see what children are doing). Which one should I cross out? Someone chose (#5). Watch while I draw it in up here. (Draw #5). Is (#5) the best piece? Why not? That's right. It's pointing in the wrong direction. It doesn't finish the design. Which one is the best piece? (Choose a child) Watch while I draw it in. Let's check to see if you're right. (Show answer slide) Yes, #3 is right--it finishes the design. The black part is toward the middle and the petal points away from the middle like all the other ones do. (Make chalk lines illustrating direction of petals; change slide)

**DESIGN 6. (wavy horizontal line design, slides #11, 12). You don't have this puzzle in your books. Look up here and find the best piece to finish this puzzle. Which piece should I cross out up here? (call on child) Let's see if you're right. (Change to answer slide) Good. (Change slide) Turn to page 3.

DESIGN 7. (wavy horizontal and straight vertical line design, slides #13, 14, page 3). Look at this puzzle. Cross out the best piece to finish this puzzle. Which one should I cross out? (Choose a child) (Child's name) thinks this one (#5) is the best piece. Does everyone agree? Good. See, we have to look for two things--the wavy lines (point with chalk) and the straight lines (point with chalk). What's wrong with (#3)? I'll draw it in. (Elicit from children why it doesn't complete the pattern). This piece (#1) finishes the puzzle. (Show answer slide). You're really good at these puzzles. Turn to the next page. (Change slide)

DESIGN 8. (Converging lines on a gradated background, slides #15, 16, page 4). Look carefully at each of the six choices to find the best piece. Cross it out in your book. Which one should I cross out? (Try to find someone who has chosen #5 to call on for the answer). Watch while I draw in (#5). Does that piece look right? Why not? Yes, these lines (point) are straight up and down. But they should slant. Now who knows which one is the best piece? (Call on a child) (Child's name) thinks (#2) is right. Let's check. (Show answer slide) Good. You must look carefully at each puzzle and at each of the six pieces before you choose one. If you do these puzzles too fast, it's easy to make mistakes. See, we have to look for the slanty lines (Change slide) Turn to next page in your books.

DESIGN 9. (Curvey bar design, slides #17, 18, page 5). This puzzle is tricky. Piece #4 looks like the best answer. Raise your hand if you think #4 is the piece. I'll draw it in up here (make drawing in empty space). Does it look right? No, the curving bars are too skinny and too close together. Which piece is the correct one? (Call on a child) (Child's name) says (#3). I'll draw it in (draw it in). That looks right to me. Let's check it. (Erase; show answer slide). Good. See piece #3 is the only piece that fits the space and finishes the design. Piece #1 has straight bars so it's no good (point to #1). The colors are wrong in piece #2. Turn to next page.

DESIGN 10. (crossed line design, slides #19, 20, page 6). Here's another puzzle with a piece missing. Which piece (point to 6 choices) is the best answer. Draw in the missing piece. Cross it out in your books. Tell me which piece I should cross out. (Choose a child, then put an X on piece child chooses) (Child's name) thinks #4 is correct. How many people agree? Good. Piece #4 has the black X that completes the puzzle. I'll draw it in up here. (Draw in the X; erase; show answer slide) What's wrong with piece #3? (Call on a child and quickly elicit answer). Yes, piece #3 has a white cross on a black background; the right piece must have a black cross on a white background. (Show answer slide). Turn to page 7.

DESIGN 11. (crossed line design, slides #21, 22, page 7). Let's do another puzzle. Only one of the pieces below will finish the design. Draw in the missing piece. (Pause while children draw.) Who can tell me which one is the best piece? (Child's name) says (#6). Watch while I draw it in. Good. (Erase and show answer slide). The missing piece must have two lines through the box this way (trace line with finger). What's wrong with piece #2? That's right, it only has one line through the box instead of two (point) and two here (point) instead of one. It's important that you look carefully at these puzzles before you choose an answer. (Change slide). Turn the page.

DESIGN 12. (zig zag line design, slides #23, 24, page 8). Look at this design carefully and draw in the missing parts in your books. You must remember two things are happening so be careful. Then cross out the best answer below. Who will tell me which piece I should cross out? (Child's name) says piece ( ) is the best. I'll draw it in up here. Does everyone agree? Let's check it. (Show answer slide) Good. Notice that the zig zag gets fatter and fatter as it goes up and down and notice the two lines coming down. Why is piece #5 incorrect? (point to #5). The fat part of the puzzle is correct but there are no lines coming down. So piece #1 is the only one that completes the design. Good. Turn to the next page. (Change slide).

DESIGN 13. (ice cream cone with a part missing, slides #25, 26, page 9). It looks like someone took a bite out of this ice cream cone. Draw in the missing piece. Cross out the piece down below that looks just like what you drew in. Who can tell me which piece to cross out up here? I'll draw it in up here. (E draws in child's choice). (Try to call on children who have made an incorrect choice and make chalk drawing on the board to illustrate why it doesn't complete the pattern.) Does it finish the design? Okay, let's check. (Change slide to answer slide). Very good! Turn your page.

DESIGN 14. (slides #27, 28, page 10) and

DESIGN 15. (slides #29, 30, page 11). Ice cream cone. Use same procedure as for DESIGN 13.

**DESIGN 16. (triangle, slide #31) Here's a triangle. Notice it has three angles or points (outline with chalk) and three sides (outline with chalk). (change slide).

DESIGN 17. (triangle with missing angle, slides #32, 33) page 12). Everyone should be on page 12. This triangle has a piece missing over here (point). Draw in the outline of the missing piece. (Wait a few seconds). See, this line goes down to the point and this line goes across to the point. (Draw chalk lines as you talk). Does everyone's drawing look like this? Now put an X on the one below that looks like what you drew. Who's going to tell me which one I should cross out up here?

Does everyone agree? Good. Let's see if #6 is right. (Change to answer slide) (Regardless of choice offered by students, E draws it on the blackboard, and asks whether it completes the pattern.)

DESIGN 18. (upside-down triangle with part of an angle missing, slides #34, 35, page 13). Draw the outline that will finish this triangle. Cross out the piece down below that looks like what you drew in. Who can tell me which piece I should cross out up here? (Call on a child) Does everyone agree? Let's check your answer. (Show answer slide) (Again, regardless of choice offered by students, E draws it on the blackboard.) So you see piece #3 is the only piece that completes the design.

**DESIGN 19. (slides #36, 37) and

**DESIGN 20. (slides #38, 39). (arrows) (These two designs do not appear in the training booklets and can be done quickly.) Look up here. This design is not in your books. Which piece should I cross out? (E draws in responses given) Does everyone agree? Let's see if you're right. (Show answer slide) Good. Notice all the arrows have to point this way. (Change slide)

DESIGN 21. (arrows to center, slides #40, 41, page 14) and

DESIGN 22. (arrows away from center, slides #42, 43, page 15). (Proceed in the following manner:

1. Instruct children to draw in missing arrow and cross out choice below.

2. Try to call on child who has made incorrect choice and draw in arrow on the blackboard.
3. Demonstrate with chalk lines the direction of the arrows given (toward the middle in Design 19 and away from the middle in Design 20).
4. Call on child for correct answer.
5. Show answer slide, then change slide to next design.)

DESIGN 23. (Arrows enclosed in a square, slides #44, 45, page 16.) Everyone should be on page 16. Look at this puzzle, then draw the arrow in the direction it should go in your books. Next draw the box around it. Cross out the best choice below. Who can tell me which one I should cross out up here? (Accept any choice given, then draw it on the chalkboard in the problem space so children can see why some choices are incorrect.) (Show answer slide as a final check.) You can see #2 is the only piece that completes the design. Turn to the next page.

DESIGN 24. (triangle away from the center, slides #46, 47, page 17). Let's look at this puzzle. Cross out the piece down here (point to 6 choices) that will finish up the puzzle. Which piece should I cross out up here? (Call on a child) (Child's name) says piece (#1) looks like the best piece. Watch, I'll draw it in. Does piece (#1) finish the puzzle? Yes. Each triangle in the puzzle is pointing away from the middle. The triangle in piece #1 points away from the middle (illustrate with chalk lines). What's wrong with piece #3? (Erase; draw piece in the empty space). The triangle points this way (point), we need a triangle that points this way (point) to complete the design (erase). So piece #1 (show answer slide) is the best answer. (Change slide) Turn to page 18.

DESIGN 25. (polygon orientation problem, slides #48, 49, page 18). Here's a fun puzzle to do. Cross out the best answer. (point to choices) Who will tell me which one to cross out up here? Someone says piece (#1) is right. I'll draw it in up here. (draw) Does that look right? No. The shorter end should point away from the center not toward the center. The best piece is # 3 . (let class answer together) Yes, piece #3 (show answer slide) is the best one. You're doing very well.

DESIGN 26. (crescent orientation problem, slides #50, 51, page 19). Here's a fun puzzle for detectives. Cross out the best answer. Which piece should I cross out up here? (call on a child) (Child's name) says (#1) is the answer. I'll draw it in the empty space. Does that look right? Yes, there are two speckled bananas here (point) and two black bananas here (point). The curved part of each banana points toward the middle so the curved part of the banana in this piece must also point toward the middle. (Change slide)

You've been working very hard today; you did very well. Close your books, make sure your names are on the front. I'll be back tomorrow and we'll finish the book. (Collect all booklets).

END OF TRAINING SESSION 1.

BEGINNING OF TRAINING SESSION 2.

**DESIGN 27. (Row of circles and triangles, slide #52). Yesterday we solved some fun puzzles together; today we'll do some more. This puzzle is not in your books. Look up here. Let's name these shapes together. (Point to shapes, and have children name each of the four shapes asking for responses for four imaginary figures beyond last triangle.) You knew what would come next (point to empty space past last triangle) because you saw how the shapes were changing and you knew the pattern. (Change slide)

**DESIGN 28. (column of squares and circles, slide #53). Here's another puzzle that is not in your books. Let's name these shapes together. (Point to and have children name each of the four shapes asking responses for four imaginary figures beyond the last circle.) You knew what would come next because you saw how the shapes were changing.

**DESIGN 29. (circles, slide #54). (Have children name each circle as you point, continuing to name four imaginary figures past last circle on the slide). All these shapes are circles. Anything you add has to be a circle. These shapes are all the same. (Change slide)

DESIGN 30. (Circles filling in, slides #55, 56, page 20). Open your books to page 20. Let's figure out what shape should go here. (Point to empty space to the right of the fourth circle) Let's name the shapes together. (Point to and name the four circles present on the slide, then point to empty space which children will label "circle"). Good. It's a circle because all the other shapes are circles. Draw the circle in your books. (Draw circle on board). Now, let's figure out what to put inside the circle. No black inside (point), a little black inside (point), more black inside (point), almost all black inside (point). Who can tell me how much black to put inside the last circle? (Child's name) says all black. (Fill in circle). Anybody have any different idea? Let's see if we're right. (Change to answer slide). Good. The circles are filling up with black. This one (point to last circle) has to be all black. (Change slide)

DESIGN 31. (two rows--squares and circles, slides #57, 58, page 21). Let's name the shapes across the top row together (point and have children name each one). Now, let's name the shapes in the bottom row. What shape should I draw here? (Point to empty space) Good. It's a circle. (Draw it in) Because all of the shapes in the bottom row are circles. Draw the circle in your books. Now let's look from top to bottom. (Point to first column) Square changes to circle but the stripes stay the same. (Point to second column) Square changes to circle but the dots stay the same. (Point to the third column)

Square changes to circle but half black stays the same. (Point to fourth column) Square changes to circle. The top one is all black, so the bottom one has to be what? (ask child) Right. All black. Now we know that our circle has to be all black (fill in the circle). Let's check to see if we're right. (Change to answer slide) Good. The missing piece had to be an all black circle because two different things happened when we went from top to bottom. The square always changes to a circle, but the inside always stayed the same. (Change slide)

DESIGN 32. (slides #59, and 60, page 22). This is like the last puzzle only this time there are only three shapes in the top row and two in the bottom row. Who can tell me what shape belongs here? (Choose a child) Why do you think it should be a circle? Good. Square changes to circle here (point to left hand column). Square changes to circle in this (middle) column, so square must change to circle in this column. This has to be a circle. Draw the circle in your books. Draw what has to go inside the circle. (Wait while children draw) Does everyone's drawing look like this? (Change to answer slide) A circle that is half black. Good. Let's check how we got this answer. Square changes to circle; stripes stay the same. Square changes to circle; dots stay the same. Square changes to circle; this top one is half black, so the bottom one has to be all black. Our rule is: whatever changes from top to bottom in this column (point), and this column (point) must change in this column. Whatever stays the same (stripes) when we go from top to bottom in this column (point to left hand column) must stay the same in this one (dots stay the same) (point to middle column) and has to stay the same in this column, so this one has to be half black (point to right hand column). (change slide)

DESIGN 33. (two by two design, slides #61, 62, page 23). This puzzle has only two shapes in the top row and one in the bottom row. You figure it out the same way we did the other puzzles. Draw in the shape of the missing piece first, then what should go inside. Who can tell me what shape I should draw in here? (E draws it) What goes inside? (E draws it) Now we can check each puzzle two ways. (Point to each column successively) We can sing a little tune. Square changes to circle (point), so square here must change to circle. Stripes stay the same here (point), so dots up here must stay the same down here. We can also check our answer working this way and let's sing our tune: (point horizontally across the top, then bottom rows) Square stays square, so circle here must be a circle here. Stripes change to dots, so stripes here must become dots here. (Show answer slide; then change slide)

DESIGN 34 (two by two design, slides #63, 64, page 24). This puzzle is just like the one we did. In the empty box here (point) draw the shape of the missing piece, then draw its insides in your book. (Pause) Who can tell me what I should draw in here? Does everyone agree? (Look for other solutions) Let's see if we're right (if no disagreements). (Show answer slide) Good. Shape doesn't change across the top row (square stays square) so shape doesn't change across the bottom row (circle stays circle). Insides change across the top row so they must change across the bottom row in the same way. How many people drew a circle with dots inside? (Repeat refrain from previous design checking columns) Who can sing our song for the other way to check our answer? (Get child to sing, with help from E, the solution by rows)

DESIGN 35. (two by two design, slides #65, 66, page 25). Look at this puzzle. Draw in the shape of the missing piece then draw what must go inside. (Circulate quickly to get an idea of what children are drawing) Who can tell me the shape of the missing piece? (Try to call on someone who has drawn in a square. Draw it on the board) Let's sing our song to check your answer. Square changes to triangle here (point) so square must change to triangle here (point). What shape must go here? (Point to where square was drawn in) A triangle (erase and draw in a triangle). Shape changes across the rows. What should I draw inside the triangle? (Child's name) says stripes. Let's sing our tune to check his answer. Half black stays half black from here (left) to here (right) (point) so stripes here must stay stripes here (show answer slide). The insides stay the same across the rows. The missing piece must be a triangle with stripes inside.

DESIGN 36. (two by two design, slides #67, 68, page 26). Here's another puzzle with a piece missing. First draw the shape of the missing piece, then draw what goes inside. Look at the six choices below and cross out the best piece. Who will tell me which one to cross out up here? (Accept any choice given and draw it in) Let's sing our tune to check your answer. Look from top to bottom. Diamond stays diamond here (point) so square up here must stay square down here (point). Stripes change to all black here (point) so stripes up here must change to all black down here (point). How many of you crossed out #4? Let's check our answer (show answer slide). Yes, #4 is the only one down here (point to choices) that completes the pattern above (change slide).

DESIGNS 37 to 42. (slides 69-78, pages 27-32). (The next series of training designs are two by two designs where two figures are in the top row and one is in the bottom row with the lower right design missing. For each design, say:) Look at the design and draw in the shape of the missing figure first. Then draw in what the inside must look like. Remember, think of what changes and what stays the same in the row or column, because the same thing--shape or inside--must change or stay the same in the other row or column. After you draw in the missing figure, cross out the correct choice below. Which piece should I cross out? (Draw it on black board) Do you all agree? (If they agree:) Why do you think it is right? (if they don't agree:) Why do you think it is wrong? Show me your choice. (Repeat procedure) (Make sure that you have the group verbalize the rule--sing the tune--for what changes and stays the same for designs #36, 38, 40. If the shape changes, give the rule for "changes" first; if it stays the same then the rule for "same" comes first. The rule should also be verbalized in each of the designs (#35-40) by a child, especially when children are making incorrect choices. This can be done very quickly without becoming monotonous by singing "our tune.") In the top row (name shape #1) changes to or stays a (name shape #2). In the top row (name inside of shape #1) change to or stay (name inside of shape #2) so in the bottom row (name inside of shape #1) must change to or stay (name inside of shape #2). (Alternate checking problem solutions using rows, then columns. Show answer slide as a final check)

DESIGN 43. (addition, slides #79, 80, page 33). In this puzzle you don't have to draw in the missing piece. Figure out the shape of the missing piece, then the insides in your head. Cross out the piece down below which is like the design you are thinking of in your head. Which one is the best answer? (E accepts any choice and draws it on the board) Let's check your answer by looking from top to bottom. Star stays star here (point) so diamond must stay diamond here (point). Now let's look at the insides. Blank inside changes to a dot inside here (point) so blank inside changes to a dot inside here (point). We're adding a dot when we go from top to bottom. Let's do the other way to make sure we're right. (Check by rows) (Show answer slide)

DESIGN 44. (addition, multiple insides, slides #81, 82, page 34). (This design is easy and can be done quickly) Look at the puzzle. Now look down below and find the piece that finishes this puzzle. Which one should I cross out up here. (Call on a child) Let's check to see if you're right. (Show answer slide) Good. Number 2 is the best answer. The shape was easy to figure out. All the shapes are circles. They're all the same. Let's look at the insides. The cross changes to a cross with dots here (point) so this cross (point) must change to a cross with dots here (point). We're adding dots as we go across the rows.

DESIGN 45. (double addition inside, slides #83, 84, page 35). Let's do this puzzle together. What does the outside shape have to be? Good. A circle. The shapes are all the same. What has to be inside the circle? When we look from here to here (left to right) we add a dot. So this circle (point) must have a dot inside. Now let's look from top to bottom. We add a triangle. So this circle (point) must also have a triangle inside. Which one below is the best answer. Number 6. Let's check (show answer slide). Good. It's a circle with a triangle and a dot inside.

DESIGNS #46 and 47.(slides 85-88, pages 36, 37, two by two designs) (Use same procedures as with earlier two by twos except children do not have to draw in the missing piece. In reviewing correct solution with the alternative changes rule, demonstrate how insides are taken away going across rows)

DESIGN 48 (slides #89, 90, page 38 and

DESIGN 49 (slides #91, 92, page 39).

(The children will be able to solve these puzzles without any introduction. When reviewing solution, point out that they are completing a circle but must also be aware of how the insides or backgrounds are changing)

3 X 3 Design Training

What we had found was that children who did solve the 2 X 2 problems with ease by the close of the training period did not spontaneously see that a similar solution process could be applied to the 3 X 3 format which had not appeared in the context of the training. Thus, there was very little transfer of the training on the 2 X 2 to the 3 X 3 problems, especially among the slow learning population. Training on three principles appropriate to 3 X 3 problems attempts to direct the child's attention to the fact that these problems are no different than those he has already mastered and that he can in fact solve them with some focused effort.

In addition, 3 X 3 training designs are included to help the student see the pertinence of the training on the double classification problems to another type of problem, namely 3 X 3 problems. Also, since three principles of solution are included in the 3 X 3 set of training designs, the child comes to understand that there are different principles that can be applied in solving these 3 X 3 types of problems and comes to see that he must identify the principle that works for the particular problem. He also learns to check whether his initial hypothesis about the appropriate principle works for the problem.

The problems are presented to help the person see that what he learned in solving the double classification problems is pertinent to a different problem format. He comes to see that he must choose a principle, test its applicability, and then check whether it works by solving the same problem in an alternate manner.

The 3 X 3 designs are presented only to students 9 years of age (CA) and over. This training procedure is designed to teach problem-solving strategies to problems found in Sets C, D, E of Raven's Progressive Matrices (1948). The 3 X 3 designs train to the following principles of solution: odd man out, addition, subtraction.

For all problems, the T should ask the children to look at all the pictures or figures in the design, and figure out what the blank one must look like and to cross it out below. Then T asks the child to indicate his/her choice for the T to cross out. T crosses out the choice and, when necessary, draws it in the blank space on the blackboard. T asks whether the other students agree with the choice, and if not, why not. If all the children agree with the choice (and it is correct), T asks the child how he/she picked the choice, and seeks to confirm the correctness of the choice by checking the application of the principle in the columns and rows, or vice versa.

DESIGN 50. (slides #93, 94, page 40, odd man out). In this picture you see there are three rows (point out) of pictures and three columns (point). (Starting with the top left figure and moving across the top row, the teacher along with the students calls out the names of the figures: House, sun, triangle; then row 2: sun, triangle, house; then row 3: triangle, house, blank.) What must go there? (Teacher should review the solution by using the usual rule:) What happens here (point to the first row) must happen here (second row) and must also happen here (third row). In the first row there are three things, a house, sun, and triangle. There are three things in the second row too, sun, triangle, and a house. The third row must also have three different pictures. You can see that we get the same choice when we go down the columns.

Here's another way you can figure out the answer to the same puzzle. Let's start at the same place (upper left hand figure), and name the figures as I point to them. (Point to each picture, left, center, and right hand columns in that order). What picture has to go in here?

DESIGN 51. (slides #95, 96, page 41, odd man out). In this picture there are also three rows (point) and three columns of pictures (point). Look at all the pictures. Who can tell me which piece I should cross out? (Call on a child). (Child's name) says piece #3. How many people agree? Good. (Show answer slide). Going across the first row we have (invite children to name each picture in turn) chair, cake, glasses.

In the second row (point) there is a cake, glasses and a chair. In the third row there is glasses, chair, and we need what? Yes, a cake. In these problems we can also solve the puzzle by looking at the columns and doing the puzzle the same way. (Quickly name each picture starting in the upper left corner). Turn to next page.

DESIGN 52. (slides #97, 98, page 42, odd man out). This design is different from the others but you figure it out the same way. Look at all the pictures in the design. Figure out what the blank one must look like, and cross it out below. Who will tell me which piece to cross out up here? (Child's name) says piece #4 looks right. I'll draw it in (draw outside shape). Does everyone else agree? (If yes), let's name the shapes together: (top row) circle, square, triangle, (second row) circle, square, triangle, (third row) circle, square, _______. Yes, triangle is right. Piece #4 has a fish inside. (Draw in the fish). Let's name the insides together. Starfish, fish, bug; fish, bug, starfish; bug, starfish, _______. Yes, fish. (Erase; show answer slide). Every row has three different shapes and three different things inside. (Check by looking at columns).

DESIGN 53. (slides #99, 100, page 43). This puzzle needs a different idea to get the right answer. Remember, whatever happens in the first row (point) has to happen in the second row and again in the third row. Who will tell me which piece to cross out up here? Does everyone agree?

Let's look at it together. Let's look at the top row (pointing to the top left figure). What do we have here? (pointing to top left hand figure), (child says: one flower) and here (pointing to middle top figure), (child says: one flower), and here (top right hand figure), (child says: two flowers). (Point to second row): remember whatever happens in the first row has to happen in this row also. What do we have here? (pointing to left hand figure) (Child says: one bud or leaf), and here (pointing to middle figure), (child: two buds or leaves). (Point to the third row, left hand figure). What's this? (Child: one flower and one bud or leaf), and this? (Child: one flower and one bud or leaf). And what must this be? Who will tell me which piece is best? (Child's name) says piece #3. Let's check to see if that's right. (Show answer slide). Good.

We can also pick the best picture to solve this puzzle by looking at the columns. This flower and this bud makes a flower and a bud. This flower and this bud make a flower and a bud. This flower and bud, and this flower and bud make what? (Two flowers and two buds). You're doing very well. Turn to the next page.

DESIGN 54. (slides #101, 102, page 44, addition). This puzzle uses the same idea as the last one. Look at it, think carefully, and try to find a way to get the correct answer to this puzzle. When you think you have the answer look at all the choices below and cross out the one you think is correct. Who will tell me which one to cross out? (Point out choices). (Call on a child). (Child's name) says piece #3. How many people agree? Let's check our answer by working down the column. (Point to upper left hand figure). There's one stick here, another one here, and two sticks here. What happens in the puzzle? Who can tell me? (Pause for answers). Right. We added this stick (top one) to this stick (middle one) to get this picture (bottom one). Now let's see if this idea works in this column. (Point to middle column) If we add this one (top one) to this one (middle one), do we get this one (bottom one)? Yes! Then we have to use the same idea in this column (right hand one). Who will come up and draw what this blank one must look like? (Call child) Good. Let's look at the answer slide. (Change slide) We could also check whether the idea is right by working across the rows. See if I add this one (top left) to this one (middle one), I get this one (right hand design). (Draw as you talk to illustrate an active process) The same idea works here (middle row). This one (left one) added to this one (middle) gives me the same design as this one (right hand design). See, adding this one (left one, bottom row) to this one (middle row) give me the same answer we got before.

Turn to next page.

DESIGN 55. (slides #103, 104, p. 45). (Subtraction.) This puzzle has a different idea. Try to figure out the idea of the puzzle in one column or row. Test out whether the idea works in the second row or column, and then you should be able to figure out what the missing design should look like. Then cross out the right answer down here (pointing to bottom). Go ahead. (long pause). Who will tell me which piece to cross out up here? (Child's name) says piece # ____. Let's check it. (Show answer slide). (Pointing to rows) Whatever happens in the first row has to happen in the second row and again in the third row. Let's look at the top row. (Point) This tree has four leaves (point to top left tree); this one has three leaves (point to middle tree). What is the difference between four and three? One. Okay, now this tree (point to second tree) has three leaves and this one (point to third tree) has two leaves. What is the difference between three and two? That is right. One. So, when we go from the first tree (pointing) to the second tree (pointing) we lose one leaf, and when we go from the second tree to the third tree(pointing to each one), we lose another leaf. Each time we go to the next picture we lose one leaf. Let's try the idea that each picture is losing a leaf with the trees in the second row. Does it work? Now look at the third row and use the idea to find out what the blank picture must look like. Remember whatever happened in the first row, must happen in the second row and must happen here in the third row. What must go here? (point to empty space). If the child called gives an incorrect answer, proceed with bottom row, pointing out each tree's leaves until empty space. Then see if a child can tell you the answer. The right answer is that the third picture has no leaves. (Show answer slide). Let's check by seeing if the same idea works when we go down the columns. (Start in the upper left hand corner comparing adjacent pictures, using same language as for the rows.) Turn to next page.

DESIGN 56. (slides #105, 106, p. 46). (Subtraction) Here are wheels with spokes. This puzzle uses the same idea as the last one. Try to figure out the idea of the puzzle in one column or row. Test out whether the idea works in the second row or column, and then you should be able to figure out what the missing design should look like. Then cross out the right answer down here (pointing to bottom). Go ahead. (long pause).

Who will tell me which piece to cross out up here? (Child's name) says piece # ____. Let's check it. What's the idea for the puzzle? (Call child) Okay, he(she) says it's because (insert child's idea). Let's check if it works in this row (top row). (If yes) Let's try it in this row (middle one). Draw in what the blank space should look like. (If no one gives an idea, say) Let's count the spokes together. (Start at top left column). Five spokes, four spokes, three spokes. What happens each time? Yes, we lose a spoke. (Point to middle column). Let's see if we lose a spoke in each picture in this column. Let's count them. Four spokes, three spokes, two spokes. We did. (Point to right column) How many spokes should be in the empty space? Let's count. Three spokes, two spokes, - one spoke. Let's check to make sure by using the same idea for each row. Start in top row. (Show answer slide)

ANSWER SHEET: RAVEN'S PROGRESSIVE MATRICES - SETS A, A_B, B

Name ____________________ **Date** ____________________

School ____________________ **Town** ____________________

Grade ____________________ **Teacher** ____________________

A		A_B		B	
A 1		A_B 1		B 1	
A 2		A_B 2		B 2	
A 3		A_B 3		B 3	
A 4		A_B 4		B 4	
A 5		A_B 5		B 5	
A 6		A_B 6		B 6	
A 7		A_B 7		B 7	
A 8		A_B 8		B 8	
A 9		A_B 9		B 9	
A 10		A_B10		B 10	
A 11		A_B11		B 11	
A 12		A_B12		B 12	

Totals __________ __________ __________

Total Score A, A_B, B __________

ANSWER SHEET: RAVEN'S PROGRESSIVE MATRICES - SETS C, D, E

Name ______________________________ Date ______________________________

School ____________________________ Town ______________________________

Grade _____________________________ Teacher ___________________________

C	
C 1	
C 2	
C 3	
C 4	
C 5	
C 6	
C 7	
C 8	
C 9	
C 10	
C 11	
C 12	

D	
D 1	
D 2	
D 3	
D 4	
D 5	
D 6	
D 7	
D 8	
D 9	
D 10	
D 11	
D 12	

E	
E 1	
E 2	
E 3	
E 4	
E 5	
E 6	
E 7	
E 8	
E 9	
E 10	
E 11	
E 12	

Totals ______________ ______________ ______________

Total Score C, D, E ______________

Total Score A, A_B, B, C, D, E ______________

Scoring Key

RAVEN PROGRESSIVE MATRICES — 1938

Sets A, AB ~~B~~, C, D, and E

Revised Order 1956

Dr. Raven has changed the order of presentation of seven items and rearranged the distractors in six others.

Score is number of items answered correctly.
Maximum score for each Set = 12.
Maximum total score = 60.

(a) Answers to these two items when pre-1956 booklets are used.

(b) Answers to these five items when pre-1956 booklets are used.

A		(a)
1	4	
2	5	
3	1	
4	2	
5	6	
6	3	
7	6	
8	2	
9	1	
10	3	
11	4	5
12	5	4

A_B	
1	4
2	5
3	1
4	6
5	2
6	1
7	3
8	4
9	6
10	3
11	5
12	2

B	
1	2
2	6
3	1
4	2
5	1
6	3
7	5
8	6
9	4
10	3
11	4
12	5

C	
1	8
2	2
3	3
4	8
5	7
6	4
7	5
8	1
9	7
10	6
11	1
12	2

D	
1	3
2	4
3	3
4	7
5	8
6	6
7	5
8	4
9	1
10	2
11	5
12	6

E		(b)
1	7	
2	6	
3	8	
4	2	
5	1	
6	5	
7	1	2
8	6	4
9	3	1
10	2	6
11	4	3
12	5	

Published by
THE PSYCHOLOGICAL CORPORATION
304 EAST 45th ST., NEW YORK 17, N.Y.

Printed in U.S.A.
56-119K

NORMS ON THE RAVEN PROGRESSIVE MATRICES

Percentile	Chronological Age in Years															
Points	5½	6	6½	7	7½	8	8½	9	9½	10	10½	11	11½	12	12½	13
95	19	21	23	24	25	26	28	30	32	32	33	35	36	36	36	36
90	17	20	21	22	23	24	26	28	31	31	31	34	34	34	35	35
75	15	17	18	19	20	21	23	26	28	28	29	31	32	33	33	34
50	14	15	15	16	17	18	20	22	24	24	26	28	30	31	32	33
25	12	13	14	14	15	16	17	19	21	22	22	24	25	26	27	28
10	-	12	12	13	14	14	15	16	18	20	20	21	22	22	23	24
5	-	-	-	12	12	13	14	15	16	17	17	17	18	19	20	21

APPENDIX B

SAMPLE SET INSTRUCTIONS

Before starting to work with the children, draw a replica of the Sample Answer Sheet on the blackboard; then proceed with the following instructions

Write your name here (hold up sample to illustrate; pause for a few seconds). Now tear off the top page. Put the top page next to the pages with the pictures like this. (Show side-by-side placement).

S1

Look at all the pictures on the first page. Point to the piece that is a whole car. (Check to see that students are pointing to the right picture). Look at the number above the picture you picked. Put that number next to S1 on your answer sheet. Who will tell me the number of the piece you picked? (Call on a child). Good. See number 6 goes in here next to S1 because this puzzle has an S1 here (point to sample booklet; hold it up). Number 6 is the only picture of a whole car. Let's try the next one; turn your page.

S2

Find the picture of the whole tree; look at the number above it. Write that number next to S2 on your answer sheet. Who will tell me which number I should write up here? (Call on a child; then fill in the number on the blackboard). Yes, number 4 is the only picture of the whole tree so number 4 goes next to S2. That's the answer to puzzle S2. Turn to page S3.

S3

This time find the white box (or square). What number is above it? Put that number next to S3 on your answer sheet. Who will tell me which number to fill in up here? (Call on a child). Yes, number 3 (fill in number at the blackboard) goes next to S3. The number above the white box (square) is 3; 3 is the answer to puzzle S3.

S4

Find the black square. Put its number on the answer sheet. Who will tell me the number of the black square? (Call on a child.) (Child's name) says number 2 is the answer. (Fill in #2 next to S4) so I put 2 here (point to line S4) next to S4. The black square is #2; that's the answer for S4. Turn to the next page.

S5

This page is S5. Find the black circle, then fill in its number on your answer sheet. Who will tell me the number of the black circle? (Call on a child). Yes, #5 is the answer. Did everyone fill in 5 next to S5? Good, #5 is the only number above the black circle. Turn to page S6.

S6

This time find the dot. Look at the number above the dot; then fill it in next to S6 on your answer sheet. Who will tell me the number of the black dot? (Call on a child). (Child's name) says #4. (Fill in #4 on replica of answer sheet). How many people filled in #4? Good, #4 is above the black dot.

name ______________________________

SAMPLE ANSWER SHEET-SETS

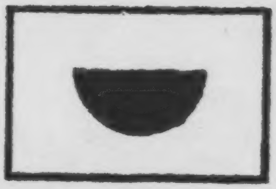
S	
S1	
S2	
S3	
S4	
S5	
S6	

5 **1**

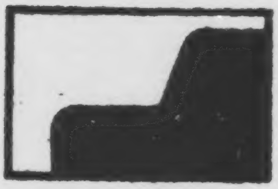
1

2

3

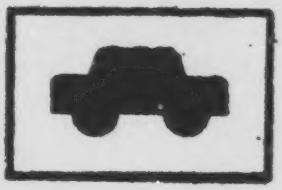


4



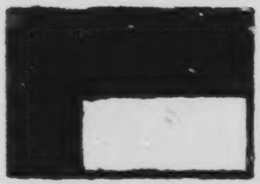
5

6

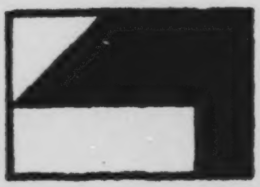


2

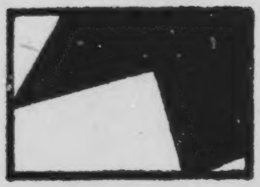
1



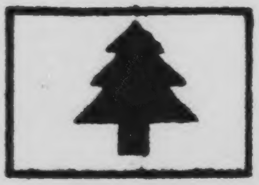
2



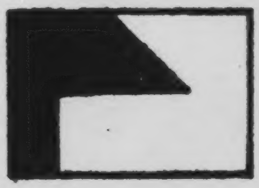
3



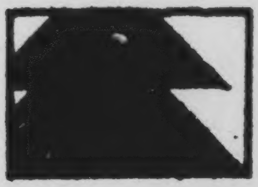
4



5

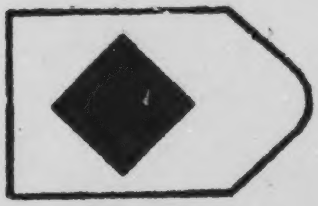


6

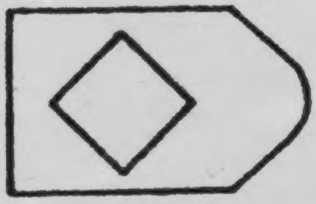


S **3**

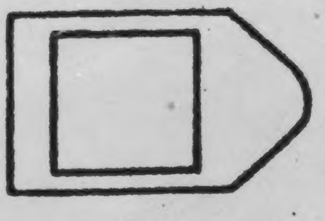
1



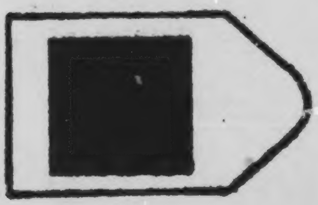
2



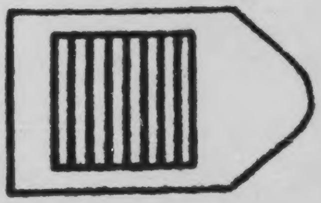
3



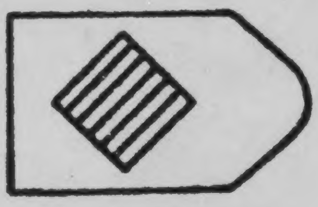
4



5

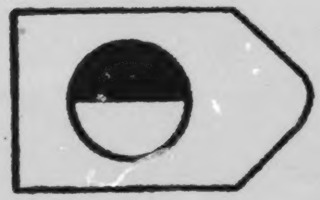


6

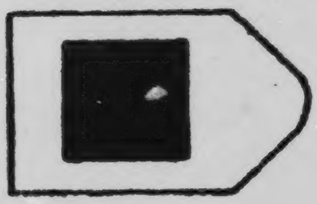


4

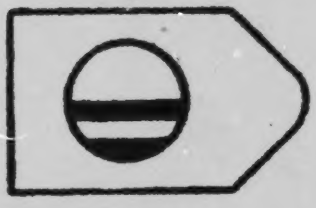
1



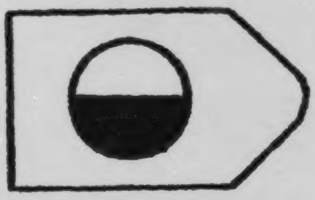
2



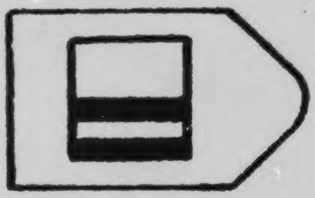
3



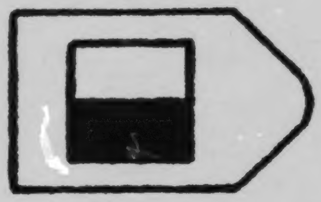
4



5

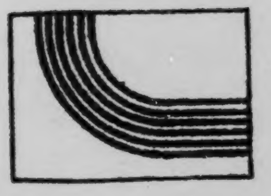


6

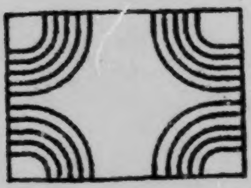


5

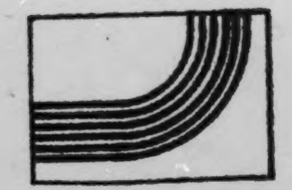
1



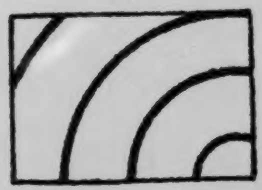
2



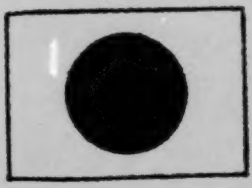
3



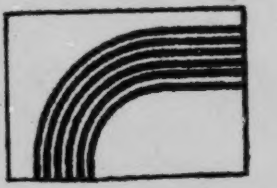
4



5

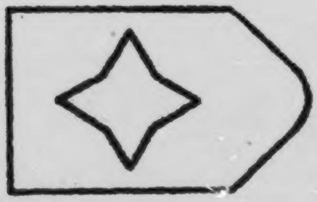


6

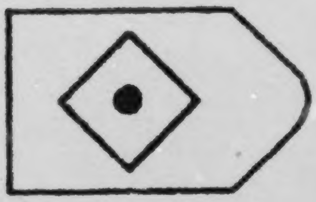


5 **6**

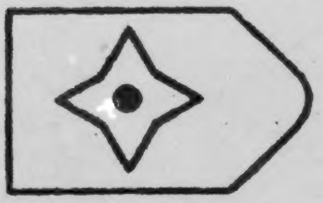
1



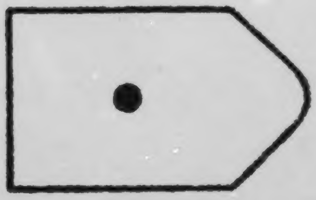
2



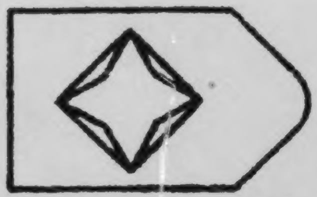
3



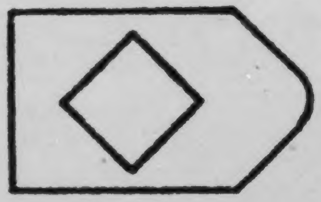
4



5



6



APPENDIX C

INTRODUCCIÓN

La Prueba Raven de Matrices Progresivas se administra como sigue:

a) Para niños de hasta 11 años de edad (incluido), administre los grupos:

A
A_B } Estos grupos están en la libreta roja
B_A

b) Para niños de 12 años de edad en adelante, administre los grupos:

A
A_B } libreta roja
B

C
D } libreta marrón
E

Direcciones para la Administración de la Prueba Raven

(para todas las edades)

Cuando los estudiantes estén sentados en el salón, déles un lápiz y una hoja de respuestas. Pideles que escriban su nombre y apellido en la parte de arriba de la hoja de respuestas.

En el pizarrón dibuje los ejemplos A.1 y C.1 (a doble tamaño approximadamente) y un ejemplo de la hoja de respuestas.

Dé a cada alumno una libreta roja y digales: "Abran sus libretas en la primera página (la página verde). Fíjense que en la parte de arriba está escrito "Set A". Ahora miren en la página de respuestas que tienen una columna "A" (muéstreles la página). En su libreta el primer problema es A.1. ¿Lo ven? A la figura de arriba le falta un pedazo para completar su diseño: "Ahora miren la parte de abajo. Tienen varias figuras, y solamente una de ellas es la que le falta al dibujo de arriba para estar completo. Vamos a ver si hallamos cuál tiene el diseño correcto. Numero 1 (enséñales con su dedo) tiene la forma correcta, pero el diseño no es correcto. Pasa lo mismo con los números 2 y 3. Los pedazos tienen la forma correcta pero el diseño no es igual al de la figura original. ¿Qué dicen Uds. del número 6? Este pedazo tiene el diseño correcto (muéstreles que el diseño es correcto) pero no está completo. Una parte es blanca.... Indique con el dedo la contestación correcta.... (pausa).... Bien, número 4 es correcta. Entonces la contestación para A.1 es 4.

"¿Entendieron lo que deben hacer? (Repita brevemente: hay una figura incompleta y hay que hallar el pedazo que le hace falta, es igual que en una rompecabezas) etc..... una vez que tenemos la respuesta la escribemos en la hoja de respuestas, aquí"... etc.). "¿Hay preguntas?"

"En cada página de la libreta hay una figura a la que le falta una parte. Uds. tienen que decir cuál de los pedazos de abajo es el que hace falta para completar la figura de arriba (Fíjense bien que solamente uno es correcto), igual que con una rompecabezas). Cuando estén seguros de cuál es el pedazo correcto, escriban el número del pedazo en la página de contestación, al lado del número de la figura."

Los problemas (o rompecabezas) son fáciles al principio, pero al final son mas difíciles. No hay trampas. Si ponen atención cómo conseguir la contestación a los problemas fáciles, encontrarán que los otros no son tan difíciles.

Vamos a hacer otro problema. Pasen a la próxima página que es de color naranja y fíjense que le falta un pedazo para estar completa. Ahora, miren cuidadosamente los seis pedazos abajo. Uno sólo es totalmente correcto. Escojan el pedazo que Uds. piensan es correcto y escriban el número del pedazo al lado del número 2 en la columna A de su página de contestación.... (pausa).... La contestación correcta es el número 5. Asegúrense que tienen el número 5 escrito al lado del número 2 en la columna A. ¿Hay preguntas?

Ahora sigan con los otros problemas. Escriban las contestaciónes en el espacio correcto en la página de contestación. Por ejemplo, cuando terminen con el problema A.12 lo siguiente es $A_B 1$ y deben escribir la contestación de ese problema en la columna que tiene A_B arriba (es decir, la segunda columna).

(Para estudiantes de 12 años de edad y mayores, digales:

"Cuando terminen con esta libreta levanten la mano, y les daré otra libreta con más problemas.")

(Para todas las edades),concluya diciendo:

"Trabajen como quieran. No hay limite de tiempo."

Nota a las maestras: Fíjense (individualmente) cómo están contestando las preguntas. Asegúrense que están escribiendo las contestaciones en los espacios correctos en la página de respuestas. Cuando el estudiante termine con el grupo B de la libreta roja (solamente a los de 12 años de edad en adelante) dele la libreta marrón con la página de contestación marcando el grupo C. Digale: "Ahora comienza con este grupo (el grupo C) y escribe la contestación al lado del número 1 de la columna C en esta página de contestación."

Diga a los estudiantes que traten de hacer lo mejor posible y que si no pueden contestar un problema, que lo dejen y sigan con los otros. Si quieren, pueden adivinar cuál es la contestación correcta.

Para la administración de la misma prueba después del adiestramiento digales: "Ustedes recuerdan estos rompecabezas. Los hicieron una vez. Ahora tendrán otra oportunidad de enseñarme lo que pueden hacer." La administración es igual que la anterior.

Procedimiento de Adiestramiento

Durante el proceso de adiestramiento el estudiante debe resolver una serie de problemas. La idea esencial es que el estudiante se dé cuenta de cómo debe enfocar los problemas, para lo cual el (la) maestro(a) le ayudará a comprender de qué se trata el problema y cómo resolverlo.

Como es necesario que el estudiante entienda el problema y encuentre la manera de resolverlo, después que Ud. le explique el tiene que tratar por su cuenta y la ayuda se le debe dar cuando no entiende o fracasa. Si el maestro(a) le resuelve todos los problemas sin que el niño se esfuerze, difícilmente entenderá cuando se lo deje por su cuenta. Por otra parte, si no se la ayuda suficientemente, el niño se sentirá frustrado, perderá interés y el adiestramiento resultará en una experiencia negative.

Más adelante se dan instrucciones específicas de cómo administrar el adiestramiento, pero siempre quedará a criterio del maestro intervenir en la manera más oportuna para que el niño (a)entienda los problemas y (b) para que tenga una experiencia agradable. Por supuesto, las intervenciones no deben violar las instrucciones especificas sino más bien implementarlas en la manera más adecuada a las circunstancias y a cada estudiante en particular.

Este procedimiento de adiestramiento ha sido usado con niños de 6 1/2 años de edad en adelante. El procedimiento completo, incluyendo los diseños 3 x 3, ha sido usado con adolescentes de 16 y 17 años de edad. Probablemente son útiles para personas mayores porque no se necesitan adaptaciones para sujetos de 17 años de edad y mayores.

Adiestramiento con estudiantes jóvenes.

Use la siguiente tabla.

Edad	Termine entrenamiento con #	Termine la sesión con #	Observaciones
menores de 7 años	35	18	
7-9 años	40	27	Pare con el # 35 si el niño no entiende.
9-11 años	49	27	No le enseñe diseños 3 x 3.
mayores de 11 años	completar la libreta (# 65)	27	El segundo día repita # 26 y 27 y siga hasta el # 65. Incluye problemas 3 x 3.

NOTA:

Usualmente se requieren 2 días o 2 sesiones dependiendo de cómo el niño coge las ideas de completar diseños.

Si el progreso es lento, sería mejor terminar la primera sesión después del problema # 18. Si el estudiante puede hacerlos en 10 minuto o menos, siga con el adiestramiento de los problemas 2 x 2[*] que comienzan con el # 20. La primera sesión usualmente termina con el # 27.

[*] Problemas: 1 x 1 ☐   2 x 2 ☐ ☐   3 x 3 ☐ ☐ ☐
☐ ☐   ☐ ☐ ☐
☐ ☐ ☐

Para sujetos mayores (EC 11-0 años) el adiestramiento sigue adelante incluyendo los problemas 3 x 3, hasta terminar la libreta. Usualmente el adiestramiento dura dos días. Primer día hasta # 27 y el segundo, repitiendo # 26 y 27 hasta el # 65.

Instrucciones Específicas

Todo lo que está subrayado es lo que Ud. debe decir.

Primera sesión de adiestramiento

E = maestro(a) S = estudiante

E tome un pedazo de papel y diga:

"Quiero que tú ves lo que estoy dibujando. Voy a dibujar algunas figuras. Mira bien lo que estoy haciendo. Verás que los dibujos cambian de cierta manera."

"Dibuja aquí lo que sigue" (indicando el espacio que sigue en el papel).

"Bien, muy bien. Aquí tenemos un pedacito, un poco más, otro poco más, más, y finalmente lo que tú dibujaste. ¿Cómo se llama lo que tú dibujaste? (un círculo, anillo, o lo que S dice.)

"Ahora, mira ésto." E dibuja lo siguiente:

"Dibuja lo que debe estar aquí."

"Lo hiciste bien." (no es necesario que S complete la figura.)

"¿Cómo se llama esto?" (enseñandole el cuadrado)

"Ahora vamos a mirar algunas figuras que ya han sido dibujadas. Abran el libro en la página 1. ¿Cómo debe ser el diseño próximo?"

I. Completando el diseño.

--- Diseño 6 - Bandera:

Con las respuestas cubiertas pregúntele:

"¿Qué es ésto?" "¿Qué hace falta?" ... Quite la cubierta.

"Enséñenme el pedazo que completa este diseño."

Si S no puede indicar el pedazo correcto, cubra nuevamente las respuestas. Pidale que le indique a Ud. qué es lo que falta en la figura. Haga que el niño mueva su dedo varias veces por las rayas de la bandera y pregúntele "¿Qué son éstos?" "¿El pedazo que falta debe tener ________?" "Si, rayas." (Descubra las respuestas.) "¿Cuál de las respuestas tiene rayas como éstas?" Si éstas preguntas no ayudan a S a resolver el problema, pruebe con una de éstas variantes:

a. Dibuje las rayas de la bandera en un pedazo de papel y póngalo en el espacio correspondiente. E indica con el dedo que las rayas van en esta dirección.

Descubra las respuestas y pregunte al S, "¿Cuál de los seis completa mejor la figura?"

--- Diseño 7 - lata de sopa:

Igual a # 6 - Si S tiene dificultades, dele ayuda dibujando la lata con el dedo del S.

--- Diseño 8 - Dibuje esta figura en el papel. E debe hacer que S se fije en las esquinas de la parte baja de la figura, dando énfasis a la forma > de las esquinas. Haga ésto con una pluma en el dibujo de S. Guarde el dibujo para los problemas siguientes.

--- Diseño 9 - Cubra las respuestas abajo y diga, "Ésta figura tiene la misma forma que la que acabas de dibujar. Enséñame con el dedo cuál es el pedazo que falta. Ahora dibujálo. Bién. Descubra las respuestas. ¿Cuál de éstas seis respuestas es el pedazo que va en éste espacio?"

Opción 1. Si la respuesta es incorrecta, cubra las respuestas, ponga las figuras # 8 y # 9 una al lado de la otra y diga,- "Ahora enséñame con el dedo como debe ser el pedazo que falta? Bien, busca aquí (descubra las respuestas) el pedazo que completa la figura."

Opción 2. Si S todavía tiene dificultades, ponga la figura 9 encima del # 8 y póngalos frente a una luz brillante. Diga,- "El pedazo que falta debe verse como ésto." Enséñele con el dedo y diga que él también lo haga con su dedo. "Bien." "Ahora, dibújalo aquí en el papel y escoge la respuesta que completa la figura." Si S no pueda resolver el problema, pase al próximo.

Opción 3. Para niños de muy corta edad o niños que tienen muchas dificultades cogiendo la idea del problema, E debe dibujar en el papel lo que S dibujó con el dedo, diciendo, "Éste es el que tú dibujaste con el dedo. ¿Cuál de las respuestas de aquí (indique) es igual a tu dibujo? Si el niño quiere hacer el dibujo él mismo, deje que lo haga.

--- Diseño 10.- "Aquí tenemos una figura igual a lo que tú dibujaste pero esta dividida en pedazo. Mira aquí abajo (las respuestas) y enséñame cuál es el pedazo que le falta." Si hay dificultades con esto, cubra las respuestas. Use la opción 2 (# 9). Diga, "Dibuja con el dedo el pedazo que le falta (en el espacio vacío)... Ahora dibújalo en el papel." Descubra las respuestas y diga al S que halle el pedazo que completa la figura. También puede usar opción 3 (diseño # 9), si lo estima conveniente.

--- Diseño 11. "Aquí tenemos otra figura como la que tú dibujaste. Ésta tiene solamente el esbozo de la figura y está dividido. Mira aquí (las respuestas) y enséñame el pedazo que falta."

Use opciones 2 y 3 (# 9) si es necesario.

II. Orientación de los Elementos en una Figura.

--- Diseño 12 - Indicando con el dedo una de las figuras arriba diga,"¿Cómo se llama ésto?" Si S no puede darle un nombre, E lo da (escudo, flor, insignia) y diga, "Vamos a llamarlo una insignia. Mira, es grande arriba y termina abajo en forma de punta." Muéstrele las respuestas y pregúntele, "¿Cuál de éstas piensas tú que debe estar aquí (indicando el espacio vacío)? Si S tiene dificultades, cubra las respuestas y diga, " En qué dirección va la punta de ésta?"(insignia izquierda arriba) - "¿Y ésta?" (arriba a la derecha)

"¿Y ésta otra?" (abajo, izquierda).... "Entonces la punta siempre está abajo." Descubra las respuestas."¿Ahora cuál de éstas tiene la punta hacia abajo?"

Si S todavía tiene dificultad, tome un lápiz con punta y diga, - "Pon el lápiz en la misma dirección que va la insignia (izquierda arriba). Haz lo mismo con lapices por las otras insignias. Después pregunta S, - " En qué dirección debe ir éste lapiz (figura baja derecha). Entonces escogela (insignia) que debe estar aquí."

--- Diseño 13 - " Qué piensas tú debe estar aquí?" Si S tiene dificultad, muéstrele cada insignia y pregúntele, - " En qué dirección va ésta?"etc. Entonces " Cómo debe estar la última? Enséñame la que es correcta." Trata de no darle muchas instrucciones verbales, a menos que no entienda el problema. Trate la estrategia de los lápices si es necesario.

--- Diseño 14 - 16 - Igual a # 13.

--- Diseño 17 - "Ésta es mas difícil. Presta atencion para que no te equivoques. ¿Cuál pedazo debe estar aquí?" Si S tiene dificultad con esta, dígale que use el lápiz. (Vea las instrucciónes, # 12.)

--- Diseño 18 - Igual a # 17.

Introducción al adiestramiento en Analogías

--- Diseño 20 - E muestra las figuras una por una, empezando de arriba y pregunta, -"¿Qué es ésto? (cuadro).... Y ésto? (círculo)... y ésto, y ésto...?" pidiendo contestaciones dos figuras imaginarias más de las que están presentes (último cuadro). Ahora preguntele, - "¿Cómo sabía que éstas figuras (el circulo y cuadro imaginarios) siguen?" (Use el mismo nombre que S da a las figuras cuando está identificando figuras particulares.) Ahora muéstrele el patrón del problema diciendo, - "Primero hay un cuadro, después un circulo, después un cuadro, un círculo, un cuadro, y para continuar con el patrón se necesita un círculo, después otro cuadro.

--- Diseño 21 - Aquí tenemos otra figura. Muéstrele el espacio vacío al fin de la fila y pregúntele, - "¿Qué debe estar aquí?.... ¿Y después de eso?" Si S no contesta correctamente o espontáneamente continue como con el # 20. E muestra las figuras y S nombra las figuras - círculo, triángulo, círculo, etc., hasta tres o cuatro figuras más que las que están presentes.

--- Diseño 22 - Muestre el espacio después del último círculo y pregunte al S - "¿En éste figura, qué debe estar aquí?" Si S no contesta correctamente, pregúntele, - "¿Qué son éstos? ¿Si pongo más, qué debe ser?" O si S no entiende, enseñele cada círculo uno por uno, y pregúntele, - "¿Qué es ésto, y ésto, y ésto? etc. Todos son círculos. Cuando no hay cambios de una figura a otra, podemos decir que todas son iguales.

--- Diseño 23 -"Dibuja lo que debe estar aquí." (Enséñele el espacio después del último círculo y de a S un pedazo de papel.)

Si S dibuja la contestación correcta, E dice, - "¿Cómo sabías que tenía que ser un círculo? ¿Está bien? Todos son círculos, entonces éste tenía que ser un círculo también. ¿Cómo sabías que había tenido que ser todo negro?.... Bien, muy bien. El interior de los círculos está más negro que el anterior hasta que tú dibujaste un círculo completamente negro.

Si la contestación del niño es incorrecta, muéstrele los círculos y diga, - "¿Mira, qué es ésto (círculo), y ésto (círculo)?" etc. hasta que todos son identificados. ¿Qué tiene que ser éste? (S: un círculo) Dibújalo (en un pedazo de papel). Ahora, vamos a ver el interior de los círculos. ¿Qué pasa del círculo 1 al círculo 2? (S: más negro, más lleno o algo parecido.) ¿Del círculo 3 al círculo 4?" (S: también más negro.) y lo mismo hasta cuando llega al ultimo circulo que falta. - "¿Cómo debe verse el círculo tuyo? (S: negro entero) E: Dibújalo."

"A partir de éste problema tenemos que poner atención a dos cosas en cada figura; algo que queda igual de una figura a la otra (como los círculos) y algo que cambia de una figura a la otra. ¿Qué cambió aquí...? ¿Qué quedó igual?

--- Diseño 24 - "¿En esta figura, qué cambia y qué queda igual?" Si S contesta solamente una propiedad, pregúntele por la otra, diciendo, - "Bien, éstos quedan iguales (o cambian) y en qué (o cuáles) cambian (o quedan iguales)? Cuando S contesta con las dos propiedades, E lo repite, diciendo, - "Si, esa es correcta, la de afuera (figura o forma) cambia y por dentro (interior) queda igual.

Para niños jóvenes, puede ser necesario que Ud. deba explicarles con más detalle. E muestra la primera figura y dice, "¿Qué es ésto?" (Usa los nombres que S le da para las figuras. E puede decirlos cuando S no lo sabe.) Después de contestar: un círculo, o redonda, etc., E continúa la línea de las figuras preguntando, - "¿Qué es ésto?" para cada forma. E dice, - "El círculo cambia a triángulo, el triángulo cambia a cuadrado, etc., mostrando cada uno con el dedo. "Entonces, todas las figuras son distintas. Vamos a ver qué queda igual en cada figura. Tal vez será necesario enseñar la idea de igual en la siguiente manera. Muestre los interiores de cada figura, diciendo:

E: - ¿Qué hay aquí dentro? (mostrando el círculo)

S: - Puntos (puntitos)

E: - ¿Y aquí? (mostrando el triángulo, haga lo mismo con cada figura)

E: - ¿Qué está dentro de todas las figuras?

S: - Puntos

E: - Bien... Entonces ¿Qué es lo que queda igual?

S: - Los puntos

E: - Correcto, el interior de cada figura queda igual.

--- Diseño 25 - "¿Qué queda igual en ésta figura? y qué cambia?"

--- Diseño 26 - (2 x 2: (2) filas y (2) columnas) Cubra la fila de abajo y muestra la fila de arriba. "¿Qué cambia, y qué queda lo mismo aquí?" Cuando S contesta correctamente, cubra la fila de arriba y hágale la misma pregunta. Pregúntele por la segunda propiedad si no lo menciona. Si S tiene dificultad con una u otra fila, muéstrele cada figura en turno y trate de que el sujeto descubra y mencione cada propiedad, una por una, usando los dos elementos de la pregunta inicial como están explicados en las instrucciones pare el diseño # 24.

Ahora, cubra todas las figuras excepto las dos en la columna a la extrema izquierda y pregúntele, - "¿Que queda igual y qué cambia cuando se va de aquí (muestre el cuadrado) a aquí (muestre el círculo)? Bien, el cuadrado cambia a un círculo y ambos tienen rayas por dentro. Pase a la próxima columna y repita la pregunta, trate de que S mencione ambas propiedades como lo hizo para la primera columna. Siga exactamente de la misma manera para las próximas dos columnas, dejando las columnas descubiertas. Para la columna a la extrema derecha, pregúntele, - "¿Qué figura debe estar aquí?" mostrando el espacio vacío. "Dibújalo." Otra vez haga que S nombre ambas propiedades. Sí la respuesta es correcta dígale, - "Muy bien."

Dele un repaso del material presentado, empezando con la columna 1: "Veamos, el cuadrado cambió a un círculo, las rayas quedaron iguales. En la segunda columna el cuadrado cambió a un círculo, y los puntos quedaron iguales. En la tercera columna .... etc. etc." Comenzando con la columna a la izquierda, muéstrele el cuadrado y el círculo diciendo, - "Todo lo que cambia del cuadrado al círculo también cambia de aquí (cuadrado - segunda columna) a aquí (circulo) y de aquí a aquí (columna 3) y de aquí a aquí (columna 4) y de aquí a aquí (columna a la extrema derecha). Entonces, ésta figura debe ser un.... ¿Qué? (círculo) ¿Porqué? ¿Qué queda igual o lo mismo de aquí a aquí (columna izquierda) queda igual de aquí a aquí (columna 2) y de aquí a aquí (columna 3) y de aquí a aquí (columna 4)? Y tiene que quedar igual de aquí a aquí (columna 5). ¿Entonces, el interior de este círculo tiene que tener...? ¿Qué?

--- Diseño 27 - Cubra la fila de abajo. "¿Qué cambia y qué queda igual o lo mismo en ésta fila?" Haga que S mencione ambas propiedades (forma o figura cambia pero el interior queda igual). Ahora, cubra la fila de arriba y pregunte a S en la fila de abajo. "¿En ésta fila qué cambia y qué queda igual?" (formas cambian, y los interiores quedan igual.) Cubra todas las figuras excepto la columna izquierda y dele las instrucciones igual que con el diseño 26, dando énfasis a lo que cambia y a lo que queda igual. Comience con la columna izquierda, después muestra la columha a la derecha y siga columna por columna de la derecha a la izquierda hasta cuando llega a la segunda columna de la mano izquierda. Siempre pregúntele las mismas preguntas, y haga que nombre ambas propiedades. Diga a S, - "Dibuja la figura que falta", y después déle un repaso.

Este es el fin de la primera sesión. Para todos los niños menores de 12 años de edad, continúe con la segunda parte en otro día. Niños mayores de 12 años de edad, podrán continuar con la segunda parte si la primera parte no les toma más de veinte (20) minutos. Pare cuando el niño se cansa.

Parte II

Si el entrenamiento total no se completa en una sola sesión, entonces comience la segunda parte con un repaso de los diseños # 26 y 27, repasando con el S lo que ya hizo al fin de la sesión anterior. Cuando termine con el repaso, empiece con el # 28. Si el adiestramiento se lleva a cabo en una sola sesión, pase al diseño # 28. - "Aquí tengo otro problema que se parece al último que hicieron, pero con algunas diferencias." Dele énfasis primero al cambio de forma y pregúntele, - "¿Qué cambia de aquí a aquí? Nuestra regla dice que si el cuadrado cambia a un triángulo, el cuadrado aquí (abajo) tendrá que cambiar de la misma manera. Dibuja la figura que debe estar aquí."

"Ahora, fíjate en la parte de adentro de las figuras. Otra vez usa la regla. El cambio de aquí (fila de arriba) a aquí (última de la primera columna) tiene que cambiar aquí tambien (segunda columna). Dibuja lo que va dentro del triángulo." Si S contesta correctamente, use # 29 para asegurarle, indicando la solución correcta.

Si S tiene dificultad con # 28, use # 29 para ayudarlo con la solución. "Así debe verse cuando todo está terminado. Todos los cambios que tiene de aquí (arriba izquierda) a aquí (arriba derecha) debe tener de aquí (abajo izquierda) a aquí (abajo derecha). Es igual al problema que hicimos aquí (muéstrele # 27). Dibuja la figura que falta (digale que copie la figura de # 29). Desde éste problema en adelante, voy a darles problemas o rompecabezas, con éste pedazo faltando (muéstrele el pedazo de abajo a la derecha)."

--- Diseño 30 & 30 A - Cubra las respuestas de abajo. Dele énfasis al cambio de forma y pregúntale, - "¿Qué cambia de aquí a aquí?" (mostrando la fila de arriba). Después muéstrele la fila de abajo y pregúntele, - "¿Cómo tiene que ser ésta figura? (Mostrando el espacio vacío.) Dibújalo aquí (dele un pedazo de papel). Ahora, ¿Qué cambia de aquí (arriba a la izquierda) a aquí (abajo izquierda)? ¿Qué tiene que cambiar de aquí (arriba a la derecha) a aquí (espacio abajo a la derecha)? Dibújalo.

Si S contesta correctamente, muéstrele primero a lo largo de la fila y después hacia abajo de la columna, diciendo, - "Círculo a cuadrado, círculo a..., mitad negro a rayas, mitad negro Descubra las respuestas y diga, - "Bien, mira el rompecabezas. Busca el que dijiste que va aquí. (indica el espacio vacío) Si S no lo entiende, refiérale a su dibujo y dígale que escoja el pedazo que se parece a su dibujo.

Si S contesta incorrecto con el dibujo, dígale, - "Vamos a ver otra vez." Mostrando primero la fila de arriba izquierda a derecha y lo mismo para la fila de abajo, dígale, - "Círculo a cuadrado, círculo a ....?" (Deje que el niño conteste. Si no puede, dígale 'cuadrado'.) Recuerda lo que cambia de aquí a aquí (fila de arriba) tiene que cambiar de aquí a aquí (fila de abajo). "¿Qué cambia de aquí a aquí? (fila de arriba) Bien, el círculo cambia a un cuadrado de aquí a aquí. Entonces éste círculo tiene qu cambiar a un triángulo (fila de abajo). Ahora, dibuja el cuadrado. Enseñe al niño a compreneer que con el interior debe proceder de la misma manera. Después pase al diseño 30 A y diga, - "Aquí lo tenemos como debe verse." Repita, - "Lo que cambia de aquí a aquí tiene que cambiar de aquí a aquí, etc. Descubra las respuestas digale, - "Escoge la respuesta correcta."

--- Diseño 31 & 31 A - Igual a 30 & 30 A.

Si S entiende bien el concepto de analogía al fin del diseño 31, continúe con los diseños 31 al 49. Para los diseños 32 - 35, con la respuesta cubierta, diga al niño que dibuje la figura que falta. Dele un repaso de la solución por analogía (enseñando cambio de forma primero). Y finalmente diga al niño que escoja la contestación correcta de entre las respuestas en la parte baja de la página. Si el niño hace el trabajo bien, el niño no tiene que dibujar la respuesta. Es suficiente si S puede decirle a Ud. cómo debe verse la figura. Después descubra las respuestas y dígale a S que le enseñe con el dedo la contestación correcta. Si S no entiende el proceso de analogía, continúe con los dibujos.

--- Diseños 41 - 49 - S indica la contestación con el dedo y las respuestas quedan descubiertas. Si S no está dibujando ni construyendo verbalmente la figura que falta, por cada tercer diseño pregunte a S porqué escogió esa contestación - indicando el proceso de resolver el problema. Esta es una manera de chequeo para que el niño no sea descuidado o adivine las contestaciones.

Si la ejecución del S no es satisfactoria, y Ud. cree que no comprende el proceso de analogía, aún cuando E usa los dibujos y explicaciones, termine el adiestramiento con el diseño # 40 si S tiene más de 9 años de edad (pero menos de 11). Si el S tiene más de 11 años de edad, continúe el adiestramiento con los diseños 3 x 3, comenzando con # 50.

Adiestramiento con los Diseños 3 x 3

Los diseños 3 x 3, comenzando con # 50, se administran solamente a los Ss que tienen 11 años de edad y en adelante (EC). Si S tiene menos de 12 años de edad y tiene dificultad con los problemas de restar (# 59 y los siguientes), administre solamente los dos últimos diseños y termine el adiestramiento. El orden de las páginas en el libro de adiestramiento debe ser así 50, 51, 52, 53, 54, 57, 55, 56, 58, 59, 61, 60, 62, 63, 64, 65, 66, y 67. Las últimas dos páginas (66 y 67) son una repetición de los conceptos presentados en los diseños 26 y 27.

Ahora vamos a trabajar con unas rompecabezas diferentes.

--- Diseño 50 - E muestra la fila de arriba con el dedo, diciendo "Casa, sol, triángulo; sol triángulo, casa; triángulo, casa, ______?" El S tiene que contestar qué figura siga las dos primeras. Si el S no dice 'sol', pregúntele por la figura que falta. Si S todavía no le responda, dele un repaso de la regla que es, - "¿Qué pasa aquí (enseñele la primera fila) tiene que pasar aquí (segunda fila) y aquí (tercera fila). Ahora, en la primera fila hay tres cosas, una casa, un sol, y un triángulo. En la segunda fila hay lo mismo. Entonces la tercera debe ser lo mismo. ¿Qué figura la hará igual a las otras filas?

--- Diseños 51 y 52, adminístrelos como # 50, primero dando S el cambio de las figuras. Cuando S trabaje cómodo con el proceso enséñele que la regla también sirve para las columnas. (Recuérdele que en los problemas 2 x 2 habían cambios consistentes para las columnas, use página 26 como un ejemplo.)

--- Diseño 53 -"Éste rompecabeza es igual a los anteriores pero tienes que poner atención a la forma de la figura y lo que está dentro (el interior). Si S no contesta correctamente enséñale la solución parte por parte, determinando primero la forma correcta y después el interior.

Para los otros rompecabezas 3 x 3, indique sumar para # 54, restar para # 59, cambio de dos formas o propiedades para # 61, y multiplicación para # 64 en ésta manera:

"Éste rompecabezas es diferente. Trata de hallar la idea del rompecabezas, después escoge la respuesta o pedazo que completa la figura lo mejor posible."

Inicialmente, deje las respuestas cubiertas y deja al niño resolver los problemas como lo hizo con los diseños 2 x 2. Sin embargo, E debe poner mucha atención a la manera que el niño usa para hallar la solucion ya que muchos Ss dan énfasis a las respuestas en vez de los componentes del rompecabezas. Si S lo hace así, muéstrele que tiene que pensar en el problema (rompecabezas) y cubra las respuestas, diciendo, - "Cuando sepas cómo debe verse el pedazo que falta, quita el papel y muéstrame con el dedo la respuesta de aquí."

Si S contesta correctamente, pregúntele, - "¿Cómo sabías que era ésa?" Si S contesta incorrectamente, vea el adiestramiento más específico abajo.

--- Diseño 54 - "Este rompecabezas es diferente. Trata de conseguir la idea del rompecabezas, después escoge la respuesta que completa mejor la figura.

Si S contesta correctamente, pregúntele - "¿Cómo sabías que era ésa?"

--- Diseño 54 - Si S contesta incorrectamente, dígale, - "Podemos resolver éste rompecabezas como hicimos con los otros. Lo que pasa en la primera fila tiene que pasar también en la segunda y en la tercera filas. Mira la fila de arriba (mostrando la figura de arriba y a la izquierda)."

"¿Qué tenemos aquí? (S: una flor) y aquí? (la figura en el medio arriba). (S: una flor) y aquí (arriba a la derecha) (S: dos flores)." Enseñando la segunda fila, dígale, - "Recuerda, todo lo que pasa en la fila de arriba tiene que pasar en ésta fila tambien. ¿Qué tenemos aquí? (mostrando la figura de la izquierda) (S: una yema) y aquí (figura en el medio) (S: una yema) y aquí (la de la derecha) (S: dos yemas)."

Antes de pasar a la tercera fila, dele un repaso de las primeras dos. (La primera) - "¿Qué pasó aquí?" S debe responder indicando de manera que se entiende que la tercera figura es la suma de las primeras dos. - "¿Qué pasó en la segunda fila?" S debe responder igual. - "Ahora, recuerda bien, todo que pasa en la primera fila también pasa en la segunda y tiene que pasar en la tercera. ¿Qué tiene que pasar aquí? (Mostrando la figura a la izquierda en la tercera fila.) ¿Qué es ésto? (S: una yema y una flor), y ésto (la figura en el medio) (S: una yema y una flor). ¿Y ésto tiene que ser qué? (S:) ¿Cómo sabías eso? (S tiene que explicar el proceso que lo llevó a esta conclusión.)

Si S no indica espontáneamente la figura correcta que falta, comience con la figura de arriba a la izquierda y dígale, - "Una flor y una flor te da cuántas flores? (S:). (diga la contestación si S no la da verbalmente.)

Enseñando la segunda fila, - "Ésta yema y ésta te dan dos yemas." Enseñando la tercera fila, - "Ésta flor y yema y ésta flor y yema te dan ________?" (dos flores y dos yemas).

"También podemos resolverlo por las columnas. Ésta flor y ésta yema te dan una flor y una yema. Esta flor y ésta yema te dan una flor y una yema. Ésta flor y yema y ésta flor y yema te dan ________? (dos flores y dos yemas) Enséñame la contestación que completará el rompecabezas. Si S no puede hacerlo, digale, - "Enséñame la figura o el pedazo que tiene dos flores y dos yemas."

Si S contesta bien, pregúntele, - ¿Cómo sabías que era ésta?

Si S no puede hacerlo, enséñele la contestación correcta y dígale, - "Ves, ésta tiene dos flores y dos yemas. Pase al próximo problema.

--- Diseño 57 - Repita el mismo proceso de 56 pero ahora dígale, - "Una persona con un brazo y una persona con un brazo te dan una personas con dos brazos, etc."

--- Diseño 55 -"Este rompecabezas tiene la misma idea. Cuando sabes la idea del rompecabezas escoge la respuesta de abajo que mejor lo completa."

Si S contesta correctamente, pregúntele, - "¿Cómo sabías que era esa?"

Si S responde incorrectamente, dígale, - "Podemos hacer éste rompecabezas como hicimos los otros. Todo lo que pasa en al primera fila tiene que pasar en la segunda y también en la tercera. Mira a la fila de arriba (mostrando la figura de arriba y a la izquierda). "¿Qué se ve diferente en cada casa? (Si S dice, - "Todas son casas, o está confundido") pregúntele, - "¿Son ésta casa (izquierda arriba) y ésta casa (la del medio) exactamente iguales?" Si no responde, pregúntele lo mismo de la segunda y tercera casas. "¿Qué es diferente? Bien. Ésta casa tiene dos ventanas (la tercera) y ésta (la segunda) tiene una. Hay otras diferencias entre las casas? No, bien. La unica diferencia es las ventanas."

"¿En qué es diferente la segunda casa de la primera?"(izquierda arriba). Si no responde, pregúntele, - "¿Cuántas ventanas hay en ésta casa (primera)? y en ésta (segunda)? Bien. Ambas tienen una sola ventana." Muéstrele la tercera y dígale, - "Decimos que hay dos ventanas en ésta casa. ¿Ésta está en el piso alto y ésta está en __________? (el bajo). Bien. Miramos a ésta casa (primera). ¿Es ésta una ventana alta o baja? Bien. Entonces la diferencia entre las casas es que las ventanas están en diferente pisos. Ésta casa (primera) tiene una ventana en el piso alto, ésta casa tiene una ventana en el piso bajo, y la tercera casa tiene una ventana alta y otra baja. Es como si tomamos ésta casa (primera) y la ponemos encima de la segunda para hacer la tercera.

Mostrando la segunda fila.

"¿Qué se ve diferente entre las casas en ésta fila? Recuerda, todo lo que pasa en la primera fila tiene que pasar en ésta también." Si S no contesta, repita el proceso de arriba.

Antes de pasar a la tercera fila, E dele un repaso de las dos filas. (Mostrando la primera fila.) - "¿Qué pasó en ésta fila?" S debe contestar indicando en una manera que se entienda bien que la tercera figura es la que resulta de juntar las dos anteriores. - "¿Qué pasó en la segunda fila?" S debe contestar en la misma manera. "Ahora recuerda bien, todo que pasa en la primera fila tiene que pasar en la segunda y tiene que pasar en la tercera. ¿Qué tiene que pasar en ésta fila? Si S contesta correctamente, felicítelo y pase al próximo problema. Si S no responde o contesta incorrectamente, dígale la contestación correcta y explíquele porqué es correcta, - "Ésta casa tiene dos ventanas en el piso alto; ésta tiene dos ventanas en el piso bajo; entonces ésta casa puesta encima de ésta debe hacer una casa con dos ventanas altas y dos bajas.

También puede enseñar al niño que el problema puede ser resuelto por las columnas, empezando por la columna izquierda. Si S no puede hacerlo, enséñele la contestación correcta, diciendo, - "Mira, ésta tiene dos ventanas alta y dos bajas." Pase al próximo problema.

Restar

--- Diseño 59 - "Éste rompecabezas tiene una idea diferente. Trata de coger la idea del rompecabezas depués busca la respuesta que completa la figura lo mejor posible." Si S contesta correctamente, pregúntele, - "¿Cómo sabías que era esa?"

Si S da la contestación incorrecta o parece que no la sabe o dice "No" dígale, - "Vamos a hacerlo juntos. Podemos resolver éste rompecabezas como hicimos con los otros. (Mostrando las filas.) Todo lo que pasa en la primera fila tiene que pasar en la segunda y otra vez en la tercera. Mira la primera fila. ¿Qué es diferente entre éste árbol (arriba a la izquierda), éste árbol (el medio) y éste árbol (arriba a la derecha)? (S: Cada uno pierde una hoja o algo parecido.) Mira la segunda fila. (Repita las preguntas mostrando cada figura.) Mira la tercera fila. ¿Cómo debe verse éste árbol? ¿Cómo sabías que tenía que verse como éste?"

Enséñele que también puede resolver el problema por columnas, empezando con la columna de la izquierda.

Si S no lo entiende, comenzando con la figura de arriba a la izquierda, dígale, - "¿Qué se ve diferente en cada árbol? Bien, el número o la cantidad de hojas en cada árbol es diferente." Si S no dice espontaneamente 'la cantidad de hojas'; pregúntele, - "¿Cuántas hojas hay en cada árbol? Muy bien, hay cuatro hojas en éste árbol (primero), tres en éste (muéstrele el segundo) y dos en éste (tercero)."

"Vamos a mirarlos más cuidadosamente. Éste árbol tiene cuatro hojas, (el árbol de arriba a la izquierda), éste tiene tres hojas (él del medio). ¿Cuanto es la diferencia entre cuatro y tres? Correcta. Uno. Entonces éste árbol (segundo) tiene tres hojas y éste (tercero) tiene dos. ¿Cuánto es la diferencia entre tres y dos? Bien, uno. Entonces cuando vamos del primer árbol al segundo perdemos una hoja, y cuando vamos del segundo árbol al tercero perdemos otra hoja. Cada vez que pasamos de un arbol al otro perdemos una hoja. Haga lo mismo con la segunda fila, excepto deje a S explicar lo que pasa, si puede. (Tercera fila) Ahora mira la fila entera y recuerda que todo lo que pasa en la primera fila tiene que pasar en la segunda y también tiene que pasar en la tercera. ¿Qué debe estar aquí?" (Muéstrele el espacio vacío.) Si S no sabe la contestación siga a la tercera fila, enseñándole que cada árbol pierda una hoja hasta (que) el espacio queda vacío. Vea si S sabe ahora la contestación.

También, enséñele que puede resolver el rompecabezas por las columnas, empezando con la columna izquierda, comparando la primera figura con la segunda, etc.

Si S todavía no sabe la contestación, dígale, y explique porqué. Muestra cada par de árboles, empezando con el árbol de arriba a la izquierda y a lo largo de la fila. Deje a S contar las hojas en cada figura: 4-3-2, 3-2-1, 2-1- y 0. Ahora dígale que le enseñe a Ud. con el dedo la contestación. Después de ésto pase al próximo diseño.

--- Diseño 61 - "Éste rompecabezas tiene la misma idea de la anterior." Repita el proceso del # 59, sustituyendo la palabra 'flor' por 'hoja'.

--- Diseño 60 - "¿Qué pedazo o figura debe estar aquí?" El adiestramiento es igual al del # 59.

Cambio de dos propiedades

--- Diseño 62 - "Este rompecabezas tiene una idea diferente. ¿Qué piensas tú que debe estar aquí?" Si S contesta correctamente, pregúntele, - "¿Cómo sabías que ésa debe estar en éste espacio?" Espere una contestación. Si S le dice correctamente porqué, pase al próximo diseño.

Si S no puede explicar la contestación o no la sabe, déle la siguiente explicación. Recuerda que en los otros rompecabezas, todo que pasa en la primera fila tiene que pasar en la segunda y también en la tercera. Muéstrele el triángulo en la primera fila, a la izquierda. "¿Qué tiene éste triángulo por dentro? (nada) ¿Y éste? (triángulo segundo) (S: tiene rayas hasta la mitad) ¿Y éste?" (tiene rayas por todo, está todo rayado, etc.).

"¿Entonces, qué cambia en los triángulos cuando vas desde aquí (izquierda arriba) a aquí (medio) a aquí (derecha arriba)? (S: los triángulos se llenan con rayas) Bién, los triángulos se van llenando con rayas de éste a éste, a éste."

"Ahora, mira la segunda fila, y recuerda que todo lo que pasa en la primera fila también va a pasar en la segunda." Muéstrele el primer triángulo en la segunda fila y pregúntele, - "¿Qué cambia de éste triángulo a éste (el medio)? Este (a la izquierda) está a medio llenar con una raya en zig-zag y no tiene rayas verticales. Éste (segundo) y éste (tercero) también están a medio llenar con una raya en zig-zag y diagonales. (Enséñele que ésta fila es similar a la primera fila excepto que la primera no tiene rayas en zig-zag en diagonales en ninguna de los tres dibujos.)

"Ahora, miren la tercera fila. Éste triángulo (izquierda) esta lleno con líneas en zig-zag, y también el del medio. ¿Cómo debe verse éste? (S: lleno con rayas en zig-zag.) ¿Qué pasa con las rayas verticales (que van de arriba a abajo)? No hay ninguna en éste triángulo (tercera fila a la izquierda), el triángulo del medio está a medio llenar. ¿Y éste, cómo debe verse? (muestre el espacio vacío). Recuerde, todo lo que cambia en la primera fila tiene también que cambiar en la segunda, y también en la tercera."

Si S no entiende, explíquele otra vez, terminando con: "Tiene que ser éste, porque está lleno con rayas en zig-zag y rayas verticales de arriba hacia abajo." También, enséñele que puede resolver el problema columna por columna.

--- Diseño 63 -"Esta rompecabezas tiene la misma idea del otro." Dele las instrucciones del # 62.

--- Diseño 65 - "Éste rompecabezas es diferente." Estos diseños solo deben usarse con niños mayores. Enséñele que la segunda figura de cada fila es dos veces más grande que la primera, y la tercera es tres veces más grande que la primera.

--- Diseños 66 & 67 - Pregunte a S cuál es el pedazo (contestación) que debe estar en el espacio vacío. Si es necesario, refiérase al diseño 26

APPENDIX D

RAVEN LEARNING POTENTIAL PROCEDURE FOR THE MODERATELY AND SEVERELY RETARDED

The learning potential assessment for the moderately and severely retarded involves a training-embedded-within testing format. The test utilized is the Raven Coloured Progressive Matrices in a formboard procedure. The test has been modified in three ways. First, the items are presented in a concrete, manipulable form by mounting the patterns and choices on pieces of cardboard material. Second, the test items are arranged in order of difficulty within the factor structures found by Corman and Budoff (1973): Simple Continuous Pattern Completion, Discrete Pattern Completion, Continuity and Reconstruction of Simple and Complex Structures, and Reasoning by Analogy. The purpose of presenting the items in this order is to maximize the probability that the training effect will transfer to subsequent, similar test items. Third, the test has been modified by providing a correction and training procedure during the test for each item missed.

The intent of learning potential is to ascertain a person's ability to learn and apply problem-solving principles under optimal conditions. The active intervention of the examiner in his/her role as trainer helps the person learn to focus on the critical features of the stimulus display, helps him/her to structure an attack on the problem, and encourages the person to actively test hypotheses regarding critical features of the problem without fear of negative judgment. The person being tested is able to change his/her perception of the situation from a purely evaluative one, with the negative affect he/she attaches to such situations, to one in which he/she perceives the role of examiner as helping.

The learning potential (LP) procedure permits one to view the individual's response to training in a variety of ways: (1) the assessment procedure illustrates how much assistance (how many trials) the person requires to reach criterion performance, (2) training on a series of similar problems can show whether after some training the child can come to solve more difficult instances of the problem--can he/she solve them on initial presentation or after only one training trial?--(3) the types of prompts given, e.g., verbal, pointing, etc., to which the person responds and which facilitates learning and communication, and (4) the degree to which the examiner needs to help the person structure an attack for solving the problems.

The training procedure, while structured in terms of stimulus materials, requires some initiative and imagination from the trainer in order to reach the child who may be capable of responding with increasing competence, but only after some address to his/her distinctive style of making connections has been made.

Procedure

Materials for the formboard version of the Raven's LP procedure consist of the actual Raven Coloured Progressive Matrices designs mounted on cardboard so the child can actually manipulate the puzzle pieces. Each design has an indented space of the size to accommodate each of the six pieces which are elevated slightly for ease in manipulation and placed as they appear in the Raven booklet. The order of the puzzles differs from that used in the Raven booklet in that similar items are grouped together in ascending order of difficulty in the formboard version. The order of presentation is A1, 2, 3, 4, 5, 6, AB3, 2, 1, B1, 2, A8, AB7, B3, A7, AB4, B4, A9, 10, AB5, 10, 6, B5, AB9, 11, B6, AB8, B9, 10, A11, B7, 11, 8, AB12, A12, B12.

The procedure itself is fairly straight forward. The tester should establish a forward-moving tempo that allows the child enough time for the thoughtful solution of each problem while engaging the child's energies toward the task. There is a total of 36 items; four test trials may be given for each item if needed. Each test trial in which a child chooses an incorrect answer is followed by training. When the child has correctly completed a design or four trials (i.e., test/train, test/train, test/train, test) have been given, the trainer proceeds to the next item. The test session is discontinued after the child misses five consecutive items after a second training trial. Two half-hour sessions are needed to complete the 36-item sequence.

In addition, short breaks may have to be taken if it is observed that the child is fatigued.

The individualized testing/training format lends itself to optimizing the child's test performance. The atmosphere to be achieved is one in which a helpful examiner, rather than an inquiring one, aids the child to become a more competent problem solver. Therefore, the examiner actually demonstrates the correct answer if the child misses an item and gives a short verbal explanation regarding the appropriateness of the piece.

However, training should not be defined simply as the process of demonstrating the correct answer. Rather, the intent is to involve the child actively in solving the problem. Training really involves teaching the child (1) to attend to the design of the puzzle, (2) to look at all the options, (3) to recognize the salient features that must be present in the correct piece, and (4) to apply those principles to the same problem and to similar problems as the task proceeds. In this LP procedure, training is embedded in the testing procedure.

It has been observed by our staff members that when the difficulty level of the problem increases and/or when the visual discrimination of the correct answer from the other options increases in difficulty, many children impulsively, and sometimes randomly choose any answer. Therefore, the examiner must be unrelentingly firm in the demand that the child actively look at the puzzle and the possible answers before making a choice. We have found pointing to be a good way to focus attention. For example, the instruction to point to each piece and then to the puzzle space when looking at the six choices works very well for some children. For other children it is more constructive for the tester to do the pointing and for the child to watch, especially when the testee impulsively grabs the pieces. The examiner will have to make the clinical judgment as to how best to handle the child.

In order to be assured that the child is actively attending to the task, the examiner helps the child focus on the task in a variety of ways. The following procedure is used.

Present the formboard, shielding the six choices with the cardboard shield provided and direct the child to look at the puzzle. Wait a few seconds and then remove the shield. Begin by pointing (or having the child point) to each piece and then to the empty space in the puzzle each time saying "look at this piece," etc., for all six choices. Be a critical observer of the child's eye movements to check that he/she is actively looking. Pause a few seconds after the pointing and say, "Okay, find the best piece."

If the child makes a correct choice, the trainer verbally reinforces by saying "good." If the child makes an incorrect choice, the trainer removes the piece and inserts the correct answer saying, "This piece is the best one." He then continues with the verbal explanation given in the specific instructions. Before another test trial is given, the six choices are removed and rearranged randomly so the order is different. It has been found useful to have some children help with this rearrangement. The trainer can remove the pieces and have the child "shake them up" before replacing them below the puzzle. Before another test trial is given, make sure that the child is reminded that the pieces are in a different order.

It is hoped that the child will be tested by someone who is acquainted with his/her style (behavior and learning) but who does not have biased (low) expectations as to how the child should perform. With children who are especially difficult to handle and motivate, proper reinforcement will be critical. For example, one child seen by one of our staff members who appeared to be untestable, worked enthusiastically for 35 minutes after the trainer discovered the child would work for the reward of touching the trainer's fur coat.

Each puzzle is numbered on the back (1 to 36). The six alternatives are also numbered on the back to reflect the order in which they are to be placed for Trial 1. Care should be taken that all materials remain out of the sight of the child until they are presented. All materials should be put in order before the test session begins and the instructor should be thoroughly familiar with the procedure so that the session moves smoothly and the child's attention is kept on task.

The child's answers are recorded on the score sheet provided by marking the number written on the back of the piece for each test trial given.

SPECIFIC INSTRUCTIONS FOR SESSIONS

Design 1 (Raven A1) Demonstration Item

(Place the formboard in front of the child making sure that the six pieces below the puzzle are in the order they appear in the Raven booklet, e.g., pieces numbered 1 to 3 (left to right) in the top row, pieces numbered 4 to 6 in the bottom row. Cover up the six choices with the cardboard shield provided. Then say, "Here's a puzzle with a piece missing." (Remove paper shield, point to each of the six pieces moving from left to right across the rows). "Look at this piece and this piece and this piece," etc. "Only one piece is best." (Pause. Then insert correct piece.) "See, it looks just like the puzzle (point to the puzzle). It is green and it has lines (trace with your finger to show the pattern in the puzzle, then in the piece). All of the other pieces fit the space but they are not right. (Remove piece #4 and insert piece #1 into the space.) Does the piece fit the space? Yes. Is it the best piece? No. It doesn't look like the rest of the puzzle--it's not right. (Remove and insert piece #2). This piece fits too, but it is not right. It is green but has no lines like in the rest of the puzzle. (Remove and insert piece #6). This one is almost right but the design doesn't go all over--see the white part (point). (Insert piece #3, then #5). These fit the space, too, but they don't look like the rest of the puzzle. (Remove pieces and rearrange them randomly.) Only one piece is the best. Now, you find the best piece. Remember, they are in a different order this time. (If the child chooses the correct piece say) "good" and then go to the next item. If the child is incorrect say, "Let me show you the best piece." Remove the piece and train--insert correct piece). See how this piece looks just like the rest of the puzzle (trace the black lines with your finger) and it goes all over (to distinguish it from #6) (Remove the piece; rearrange randomly the choices). The pieces are in a different order this time. Find the best piece. (If the child chooses an incorrect piece, train again until he/she is correct or until the child has attempted the puzzle four times.)

Design 2 (Raven A2)

(Direct the child's attention to the puzzle by shielding the pieces with the cardboard.) Look! This puzzle has a piece missing. (Point to each piece and to the empty space each time saying:) Look at this piece and this piece and this piece, etc. (Pause.) You find the best piece, the one that finishes up the puzzle. (If the child is correct say "good" and go to the next design; if the child is incorrect, remove the piece saying "I'll show you the one" and train). See (point to puzzle piece), it looks just like the rest of the puzzle. Watch my finger (trace the diagonal lines through the piece to illustrate.) See how the lines go through the piece. (Remove all pieces and then rearrange randomly.) Now you find the best piece. If the child chooses the correct piece say "good" and go to the next design. If child is incorrect, train. (The child can have four attempts for getting the answer.)

Design 3 (Raven A3)

Follow this basic procedure.

1. Show formboard; block six choices with cardboard.

2. Uncover choices and point (or have child point) to each of the choices.

3. Instruct child to choose best answer.

a. If child is correct say "good" and proceed to the next design.

b. If child is not correct, train:

1. Remove piece and insert correct piece.

2. Point out relevant features of the piece saying,"See how this piece looks just like the rest of the puzzle. See how the blue lines go through the piece (trace verticals then horizontals) to fill it up."

3. Rearrange all the pieces. Then child attempts task again. Say, "Look at this piece," etc.,"and find the best one."

FOR ALL OF THE REMAINING DESIGNS (4 to 36) JUST THE VERBAL EXPLANATION IS GIVEN. THE SAME PROCEDURE IS USED THAT IS OUTLINED FOR DESIGN 3.

Design 4 (Raven A4)

Repeat procedure used for Design 3. If training is necessary say: "The best piece has dots just like the puzzle." (Illustrate by point to show the regular spacing of the dots. Begin with the column of four dots in the puzzle just to the left of the tab space).

Design 5 (Raven A5)

These flowers look (point to piece) just like all the others (point to puzzle) and there are just enough. See. (Begin with the row of three flowers just above the tab space saying as you point to each: "flower, flower, flower" then point to and label the flowers in the piece moving left to right.)

Design 6 (Raven A6)

This design is difficult and an important part of the training is to get the child to attend to each of the six alternatives and not just impulsively choose one. For training, after inserting the correct piece say, "See how this piece looks just like the puzzle. Watch my finger. Here's a black line (trace your finger across the line and through the box), a green line (trace with finger), a black line, etc. The lines here (point to piece) look just like those in the puzzle."

Design 7 (Raven AB3)

(While training for this design point to vertical lines in each piece). Straight lines go up and down. (Trace the horizontal lines in each piece). Two curved lines go across. All pieces look just the same.

Design 8 (Raven AB2)

(Insert the correct piece.) All these pieces look the same. (Start with top left piece). Two curving lines, two curving lines, etc. The curving lines always go up and down.

Design 9 (Raven AB 1)

(Stress the horizontal and vertical directions of the lines while training. Tracing the lines in each piece.) Two lines go up and down; two lines go across. You must look in two directions.

Design 10 (Raven B1)

See all the flowers look just the same (point to each one). This piece has the flower that looks just like all the others.

Design 11 (Raven B2)

All four pieces look just the same. They are the same size; each one has a blue cross (outline with finger).

Design 12 (Raven A8)

See how two lines go up and down and two lines go across the piece (trace with your finger to show how it completes the pattern) to finish up the puzzle.

Design 13 (Raven AB7)

These look like fancy arrows. See the direction in which each one points; watch my finger (outline to illustrate directionality) see how each one points--toward the middle.

Design 14 (Raven B3)

See how these two triangles (point to triangles in the top row) look just like these two (point to bottom row). These two (point to left column) point in this direction (point) and these two (point to right column) point in this direction.

Design 15 (Raven A7)

Let's see the directions that the lines go. (Begin with the vertical line at the top of the puzzle and trace with your finger). This black line goes up and down; this black line goes across the box.

Design 16 (Raven AB4)

Watch my finger as I trace the black line. (Outline the shape with finger.) This piece finishes up a square shape. Now you trace the shape with your finger.

Design 17 (Raven B4)

The black is in this part of each square--close to the middle (outline the circle with finger). What shape is this? Yes, a circle.

Design 18 (Raven A9)

Look at this piece (trace with your finger to show how the pattern of the stripes is continued through the piece.) Point to illustrate, e.g., "white stripe" (trace from top to from top to bottom), orange stripe, white stripe, etc.

Design 19 (Raven A10)

See how this piece completes the pattern. The black stripe gets fatter from here to here (point left to right).

Design 20 (Raven AB5)

This is the best piece. Watch while I trace the design. See (begin with outside circle and work to center. Point to and name the circles).

Design 21 (Raven AB10)

This piece is best. Look where the fat black line is (outline each black line starting with top left) inside near the middle. So, in this piece the black has to be inside near the middle.

Design 22 (Raven AB6)

This piece finishes up an oval shape. Watch while I trace it. Now you trace it with your finger.

Design 23 (Raven B5)

This is a funny shape. Round here (point to left column) but corners here (point to right column). Watch while I outline the shape.

Design 24 (Raven AB9)

Look where the black is (cover up right column) (point to each black area). The black is up and down on this side in the corners (remove cover) so the black is up and down on this side. The black is always in the outside corners (point).

Design 25 (Raven AB11)

See how the lines make a shape--a diamond. (Begin with outside lines.) Watch while I trace the lines. The yellow part is outside the design (point).

Design 26 (Raven B6)

Look at this piece. White corners close to the middle. (Point and outline.) Black changes to stripes here (point to top row) so black changes to stripes here. (Point to bottom row.)

Design 27 (Raven AB8)

See this funny shape. It's round on this side (outline with finger) and pointed on this side (outline with finger).

Design 28 (Raven B9)

See how things are changing (start with left column). Diamond changes to square here (point) so diamond must change to square here. Blue inside changes to stripes here (point to row) so blue inside must change to stripes here (point). You must look for changes in two directions.

Design 29 (Raven B10)

See how things are changing (start with rows) diamond changes to star (top row) so diamond must change to star here (point to shape in bottom row). Let's look at inside (point to shapes in left column). Nothing inside changes to dot inside here so nothing inside changes to dot inside here.

Design 30 (Raven A11)

(Outline the vertical lines through the piece saying:) See how these lines curve in the same way as these (point to vertical lines to the left). (Trace the horizontal lines through the box). See how these lines curve in the same way as these (point to horizontal lines just above). You must look in two directions to get this one.

Design 31 (Raven B7)

Blue triangle changes to striped triangle (point to left column, then to right column) when we look from top to bottom. The corner of each triangle points toward the middle (outline with finger).

Design 32 (Raven B11)

See how the shapes change. Diamond changes to square (point to left column) here so diamond changes to square (point to right column) here (continue with top row) cross inside changes to nothing inside here so cross inside changes to nothing inside here (point to bottom row).

Design 33 (Raven B8)

(Again, point out changes first by row then by column). Circle changes to square (point to top row) so circle must change to square here (point to bottom row). Now look up and down. Stripes change to half black here (point top to bottom) so stripes must change to half black here.

Design 34 (Raven AB12)

Let's look across the rows. Thin blue line changes to fat blue line (point to top row) so thin blue line changes to fat blue line here (point to bottom row). What changes up and down? Blue line changes to _________ here (point to left column) so blue line changes to __________ here. (point to right column).

Design 35 (Raven A12)

See how all of these lines match up to finish up this design.

Design 36 (Raven B12)

Let's look at the changes inside. (Point to left column). Diamond and circle inside change to just circle inside. From top to bottom we take away a diamond. So here (from top to bottom right column point) we lose a diamond. The best answer is a plain square.

11/73

PICTURE-SENTENCE TRAINING INSTRUCTIONS

Today, we're going to learn a new language. You'll be seeing designs and pictures. It's really fun. We'll work together today so that when I come back you'll be able to do it by yourself. First, put your name and the date, etc. on your booklet. [Indicate on blackboard how cover sheet information should be filled out.]

PAGE 1

Let's look at the first page. Look at the first row. You see a square. Look at the four pictures in the row after the square carefully. Who can tell me what the square stands for? [*E* calls on children until the idea that square □ stands for boy is established.] How did you know that the *square* stands for *boy*? [Get the idea that the boy is common to all four pictures and hence the square must refer to the boy in each picture.]

Right. The boy is in every single picture in that row and so the square must stand for boy rather than for standing up, walking, running or any of the other activities that the boy is doing in the pictures. Now, we've learned our first word in our new language. Whenever you see a square □ it will always stand for *boy.*

Let's look at the second row. Look at all the pictures carefully. Who can tell me what the half moon ⌓ stands for?

[E waits to get the association that half moon stands for dog.] How did you know that half moon stands for dog? [E elicits the idea that dog is common to all the pictures in the row: dog walks, dog is lying down, dog is jumping, all have the idea of dog doing something.] Now, we've learned the second word in our new language, that half moon always stands for dog. Whenever you see a half moon, it will always stand for dog. Okay. Now, we're going to look at a few examples to see whether you understand how our new language works. Let's look at Number 1. We see the half moon. Cross out the picture that the half moon stands for. Which one did you cross out, the picture of the boy, the dog, the monkey, or the woman? Raise your hand and tell me which one you crossed out. That's right, we cross out the picture of the dog. Does everyone agree? [Look for dissenters.]

Let's look at line Number 2. There is a square in the box next to the number. Which picture do we cross out here? Cross it out in your books, and then raise your hand to tell me. That's right, it's the fourth one, the picture of the boy because square stands for boy in our new language. [E during this interval should be looking at all the children's booklets to be certain that they're crossing out the right one.] Did everybody cross out the boy? Everybody who did, raise your hand. [This may be unnecessary.]

Let's look now at line Number 3. This time, we have a picture of a boy and we have to cross out the design which stands for that picture in our new language. Cross it out in your books and then raise your hand to tell me which one you crossed out. That's right, it has to be the third one because that is a square and square stands for boy in our new language. Does everyone agree?

Let's look at line Number 4. This time, we have a picture of a dog. Cross out in your book which design stands for the word dog in our new language. Raise your hand so you can tell me which one it is. [Leave a pause while T (Trainer) walks around the room checking the students' progress.] Who is going to tell me which one it is? Okay, (child). which one is it? That's right, it's the picture of a half moon. Let's look at line Number 5. You do this one by yourself. Cross out the design that stands for the picture in your book and then raise your hand. That's right, we cross out the square because square stands for boy. [By the time this page is done, the trainer, T, should have walked around to ascertain that all the children in the room have gotten the idea and are marking them out correctly. Take time at this point with a child who has not grasped the idea and individually show him where his mistakes are.]

PAGE 2

[To the class:] Let's all turn the page and look up at the top of the next page. You see the "words" for boy and dog that we've already learned in our new language. Now we're going to learn two new words. Look at the picture beside the triangle. Who can tell me what word the triangle stands for in our new language? [Call someone with their hand up and say:] That's right, all the pictures in the row show a woman and so the triangle must stand for woman rather than for standing up or walking or sitting or any of the other activities that the woman is doing in the pictures. The design triangle stands for woman because the woman is in every single picture. Now, we've learned another word in our new language. Triangle stands for woman. Whenever you see a triangle, it will always stand for woman.

Now, let's look at the second new word. Who can tell me what word the rectangle stands for in our new language? Raise your hand. [T calls on one of the children who indicates the cat is the correct one.] That's right. The cat is in every single picture in that row and so the rectangle must stand for cat rather than for lying down, jumping, sitting or any of the other activities that the cat is doing in the pictures. The design rectangle in this case stands for the cat because the cat is in every single picture. Now, we've learned another word in our new language. Rectangle stands for cat. Whenever you see a rectangle, it will always stand for cat.

You can look up at the top of the page whenever you're not sure that you're right. Look carefully at all the pictures that you can choose from and cross out only the one picture or design that stands for the design or picture that's next to the number. Let's look at line 6. Cross out in your book the picture that stands for triangle in our new language. Who will tell me which one they crossed? That's right. The picture of the woman, because woman stands for triangle in our new language.

Who will tell me which one they crossed out in line 7? That's right, rectangle, because that design stands for cat in our new language. [Similarly, Numbers 8, 9 and 10.] Okay. You're doing very well, and all of you are absolutely right. Let's turn to the next page which has a 3 at the bottom.

PAGE 3

At the top of the page you'll see all the words you will need to use on this page. You see that we're just going to work with the same words that we learned on the last two pages, only this time, next to the number we have two words or two pictures together and you have to cross out the box which tells what these designs or pictures stand for in our new language. Let's look at line 11 together. Cross out in your book the pictures that best explain the meaning of the two designs in the box next to number 11. Now, who can tell me which box they crossed out? It's the first one with the woman and the dog because triangle stands for woman, and half moon stands for dog in our new language and the woman and the dog are in the same order in the first box. The second box has a woman, too. Who can tell me why the second box isn't right? That's right, the second box has a boy and we're looking for a dog.

Let's look at line number 12. Pick out the designs that stand for the pictures of boy and then a woman, which are in the box next to the number. Raise your hand and tell me which one you crossed out. That's right. The box with a square, then a triangle because there is a boy, then a woman. [Similarly, number 13 and 14.]

Now, let's look at line 15 because you see that there are three designs in the box beside the number. The square comes first, then the triangle, and then the half moon. Cross out the picture that shows what the designs stand for. Which one did you cross out? That's right, it's the last one because there is a boy and then a woman and then a dog, just like the designs are in the box next to the number. Let's look at why the first box is wrong. There's a boy there, and a woman comes next, but what's wrong with this picture? [Pause.] That's right. There's a cat when the design that stands for dog is the one we have to look for.

Let's turn the page to page 4.

PAGE 4

This time we have two words that we know already at the top of the page and we're going to learn two new words.

Let's look at the upside-down V. Who can tell me what the upside-down V stands for? That's right, it's lying down. The boy is lying down, the woman is lying down, and the dog is lying down, and so on. The upside-down V stands for lying down. Whenever you see an upside-down V, it will always stand for lying down.

Let's look at the next line. Who can tell me what the slanting line stands for in our new language? That's right, it stands for walking. The dog is walking, the boy is walking, the woman is walking, etc. Whenever you see a slanting line it will always stand for walking.

As you do each one look carefully at all the pictures and designs and try to figure out what picture or design best explains what's in the box next to the number.

Okay. Let's look at line 16. Cross out the picture that tells what the designs are about. Who will tell me which one they crossed out and why? [Get the child to "read" the message in the stimulus box, i.e., boy walking.] Okay. Let's look at line 17. Find the box with the design that best describes the picture. Put an X through the box. The picture shows a dog lying down. Who can tell me what the sentence in the first box is? [Have children read the sentence out loud in each of the four choices.] Dog lying down. Who will read the second box? Boy walking. The third box? Dog walking. The fourth box? Boy lying down. Which one did you cross out? [Call for a correct answer.] Okay. Line 18. Who's going to tell us what the sentence is? [Have child read it.] Who's going to tell us which one to cross out?

Line 19. Boy lying down. Look at the pictures, then cross out the best answer. Let's read the different choices we have. [T has the children read successively in each of the boxes.] Boy walking, boy lying down, dog lying down, dog walking. Okay. Which one do we cross out?

Let's look at line 20. This one has three words in our new language. Who can read the sentence for me? It's boy and dog walking, so we have to look for boy and dog walking. Okay. Which one do you choose? Very good.

PAGE 5

Let's turn the page to page 5. You can see at the top of the page that all the words in our new language that we have learned before are up there and there are two new words.

Who can tell me what the backwards E stands for? Right. Sitting. The woman is sitting, the man is sitting, and the cat is sitting. Everyone in the row is sitting so the backwards E stands for sitting rather than woman, or man, or cat. Backwards E stands for sitting. Very good. Whenever you see a backwards E, it will always stand for sitting.

Who can tell me what the crossed lines stand for in our new language? Standing. The boy is standing, the woman is standing, the man is standing, and the dog is standing. Everyone in the row is standing, so the crossed lines stand for standing. Whenever you see crossed lines, it will always stand for standing.

Let's look at line 21. You can do the whole page by yourself and then we'll check them together. Okay. Start. Put your pencil down when you finish and don't turn the page. Remember, look carefully at all of the pictures and all of the designs before you decide which one you're going to cross out. [The pattern of review can be as for page 4. Have a child read the sentence choices in each box when the stimulus is a picture.

Then ask which choice they crossed out.]

PAGE 6

Okay, let's turn to page 6. On this page we have some of the words we learned before and two new words. Who can tell me what the upside-down T stands for? That's right. Pulling. How did you know that? The boy is pulling the dog, the dog is pulling the woman, the boy is pulling the boy, and the woman is pulling the woman. Everyone in the row of pictures is pulling something, so the upside-down T stands for pulling. Whenever you see an upside-down T, it will always stand for pulling.

Who can tell me what the half square stands for? That's right. Chasing. The dog is chasing the dog, the dog is chasing the boy, the boy is chasing the boy, the boy is chasing the dog. Whenever you see a half square, it will always stand for chasing. Does everybody understand? Let's see if we can use these new words in the examples on this page.

Look at line 26. Read what the new words in our new language stand for, and cross out the picture that tells what the new words in our new language stand for. Remember to look at all the pictures and all the symbols and say them all quietly to yourself. When you're finished, put your pencil down and check your answers. and then we'll all check our answers together.

[T does as before. He has somebody read the sentence and has somebody tell him what the answer should be and gets everybody to agree or not. With the designs and the choices, he has different children read the designs in sequence and then someone tells him which is the correct one for the pictures. Complete page 6.] All of you have a page at the end of your booklet that looks like this one and has a 9 at the bottom. [Holding up.] This page has all the words in our new language we have learned together. Pick up the booklet, turn to the last page, check that it looks like the one I'm holding up, and tear it off carefully. If you don't have one, raise your hand and I'll give you one.

PAGES 7 and 8

Now turn to page 7 and place this vocabulary page next to your work booklets so that you can look at what the words in our new language mean if you don't remember them. When you finish this page, turn the page and complete the next, and last, page. [No new items are taught and the child simply goes through the remaining items on the next two pages. Then the items are all checked collectively in class.]

Remember to place the vocabulary page next to the page you are working on. Look at line number 36. This time you see four words together. Who can read these four words out loud? Raise your hand. Now tell me which picture is the one that shows what the four words stand for? Right. No go ahead and finish the booklet.

[When all students have finished, E reviews the two pages as before and says in leaving:] You've worked hard today and are doing very well learning this new language. I'll be back another day so you can each try to use this language again.

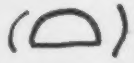
TEST INSTRUCTIONS.

(from one day to one week later; preferably the next day to minimize forgetting.)

You remember that the last time I was here we learned a new language working with pictures and different shapes. Today we are going to use the language we learned last time. First, write your name on the front page of the booklet, and in the next space, place today's date which you can copy from the blackboard. Has everyone written in his/her name and the date? Good. Turn to the first page. Now tear off the last page of the booklet and put it next to your work booklet like you did last time. (E make sure that the child's code sheet is on the side away from which he writes.)

Now look at the picture in the first line. What is that? (A cat.) Look at your code sheet and find the row with only pictures of the cat. What shape is next to those pictures? (a box, a rectangle). Now cross out the picture of the (box -- use the child's word for the figure.) Which one did you cross out? Right, the last one.

Now, let's look at the second line. This time we see two shapes and they tell us something. Look at your code sheet for the first shape. What is it? It's a square, a different kind of (box) because all the sides are the same size, not like in the first one where the sides had different sizes. Look at the code sheet and see what the square means? What does it mean? Yes, a boy. Look at the second shape, and look at the code sheet. What does that mean? Yes, a cat. What do they both say together? Yes, boy cat. Now cross out the picture which has boy cat. Which one did you cross out? Right, the third one.

Now go on and do the next one yourself. What does the picture next to the number 3 say? Right, dog running. Now look at the code sheet and pick out the shapes for dog and running and then find the box that has those shapes in it in the right order so the shapes say "Boy running". Then cross it out. (children work by themselves) Which one did you pick to cross out? (second box is correct) Call on a child. (......): Which box did you cross out? Does everyone agree with _____ _____ child's name? No? Which one did you pick? Ultimately, trainer says, See dog is this shape, a  half moon, so the first shape has to be that one because they must go in the right order. The second shape which stands for running must be a half circle, ∩ , so the right box to cross out was the second one. (Trainer draws the shapes on the blackboard to illustrate). Do you all understand what to do? First you look at the box next to the number, and you figure out what the picture or the shapes say. Then you look at the code sheet like you did last time to see which box you should cross out because it says the same thing as the box next to the number.

Now work on #4, the last one on the page. Cross out the box which says the same thing as the shapes say after you have looked at the code sheet. Remember everytime you look at a puzzle, you must look at the code sheet which will help you figure out the puzzle.

After you finish #4, go on to #5, and so on until you finish. If you have any problems, or don't understand something, just raise your hand and I will come over. Remember that what we are doing today is just the same as we did last time I was here. Take as much time as you need. Go slowly and carefully, and try to figure out every puzzle.

When you have finished, check your work to make sure your answers are completely right. Then raise your hand, and I will collect your booklet.

TRAINING BOOK

NAME ______________________ DATE ______________

GRADE __________ TEACHER ______________________

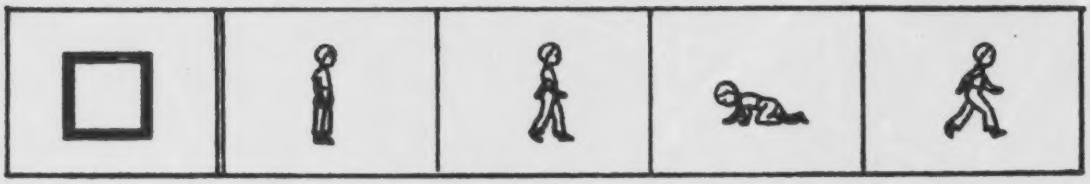
SCHOOL ______________________

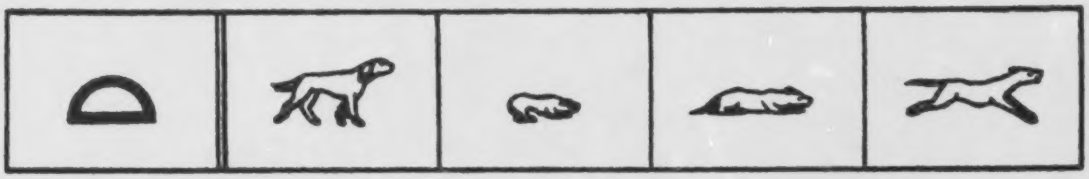
BIRTHDATE (MONTH, DAY, AND YEAR YOU WERE BORN) ____________

SEX (CHECK ONE) _____ BOY _____ GIRL

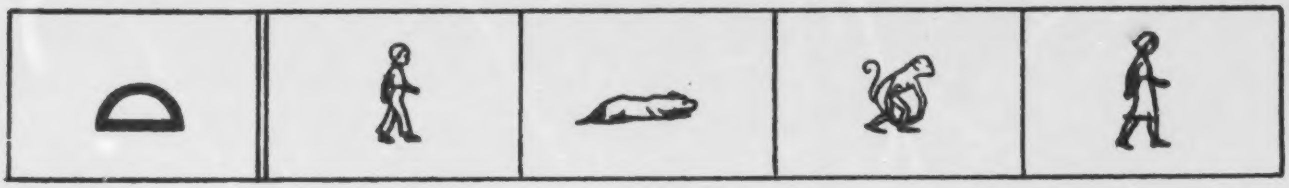
riep

THE PICTURE-WORD GAME

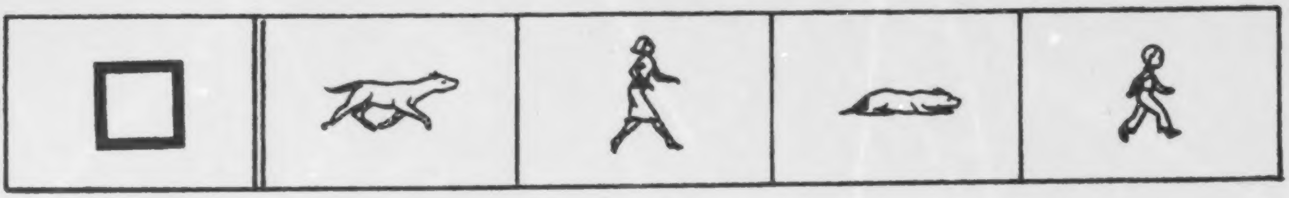




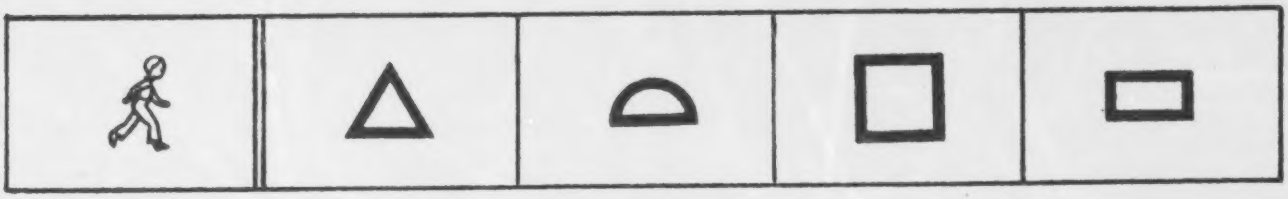
1



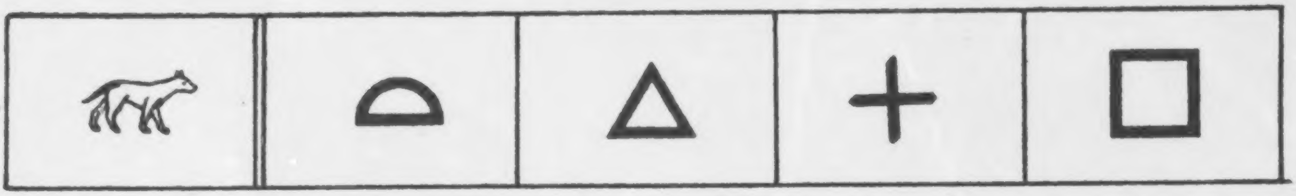
2



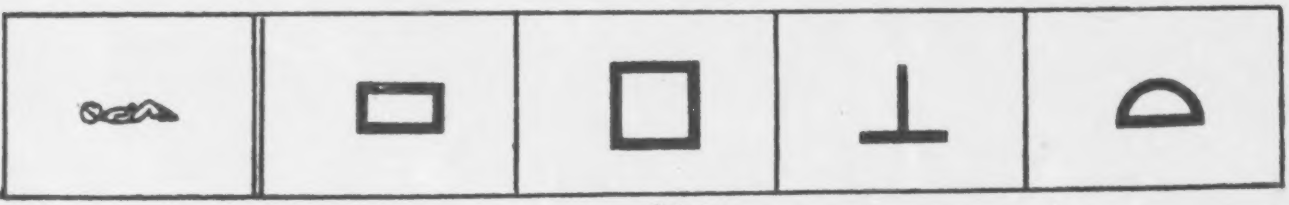
3

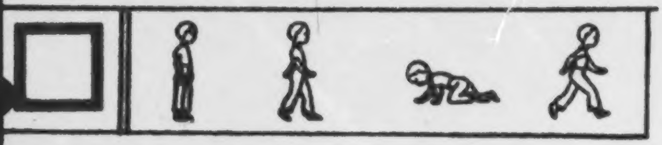


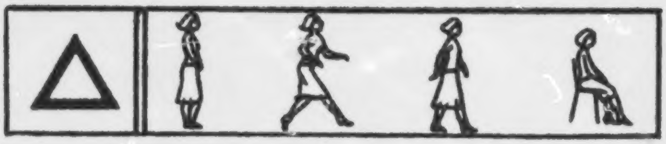
4

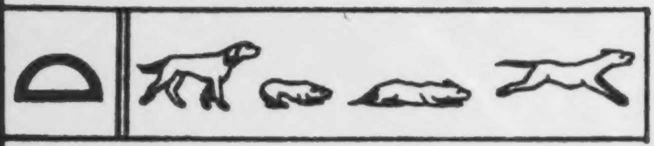


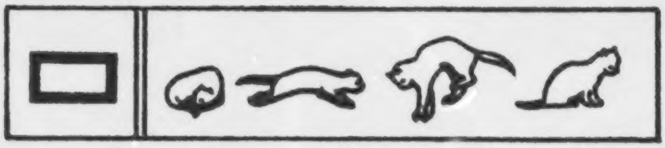
5

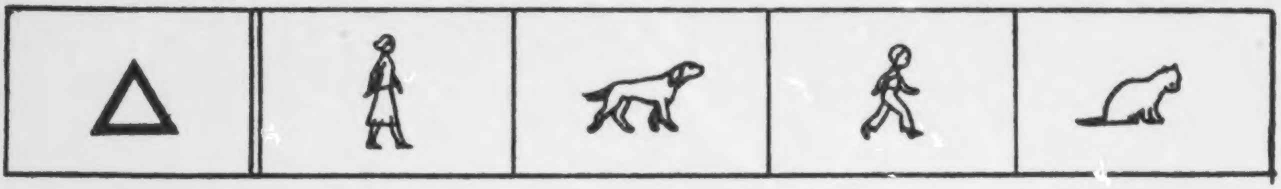


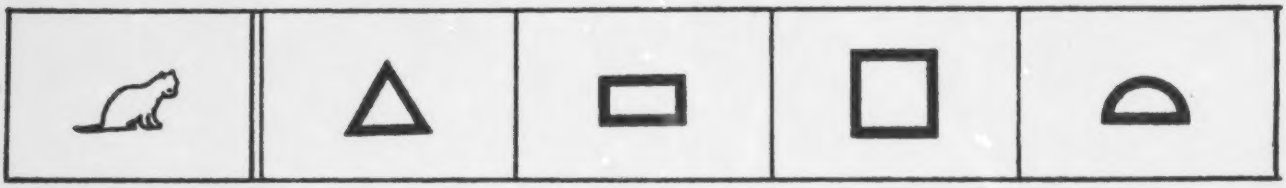


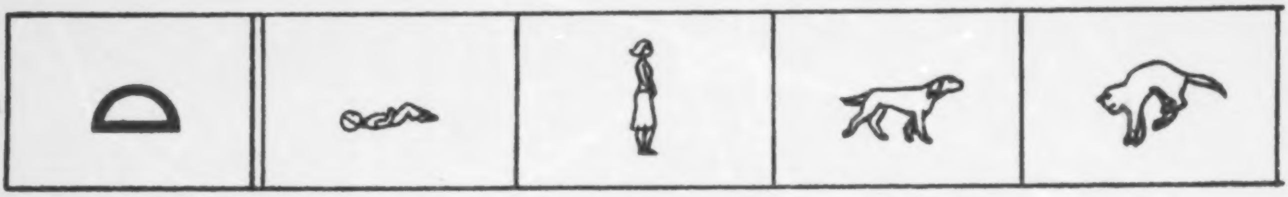


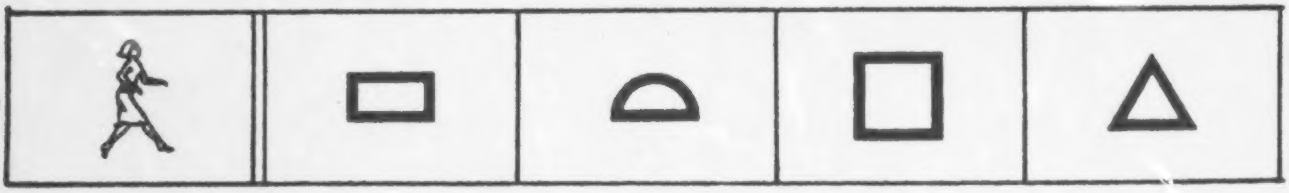


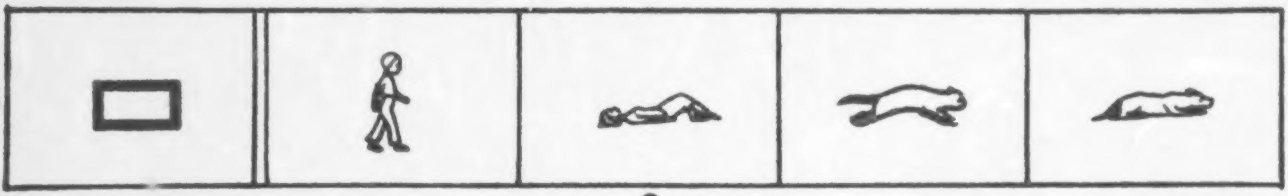


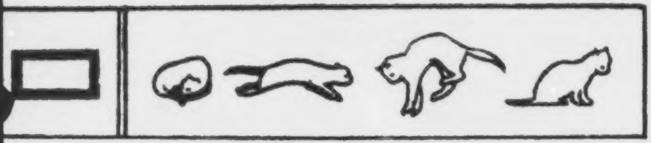
6 

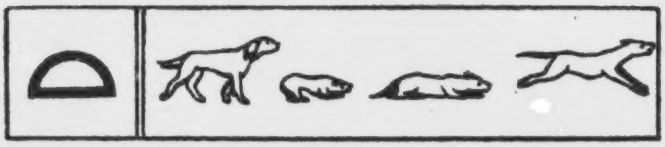
7 

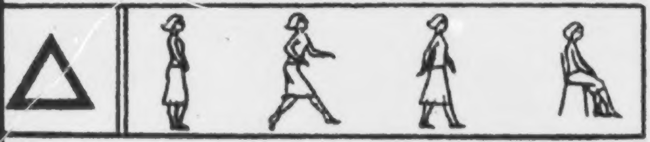
8 

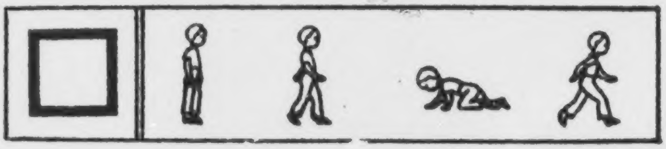
9 

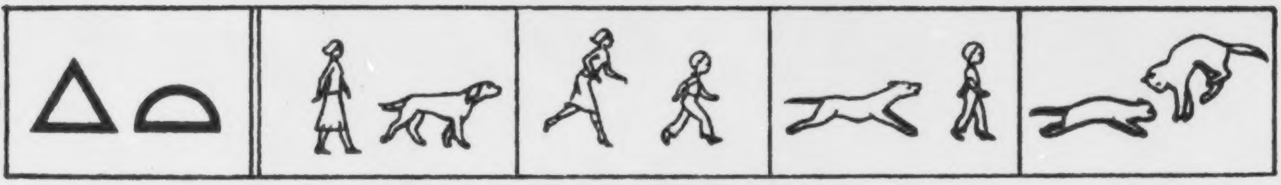
10 

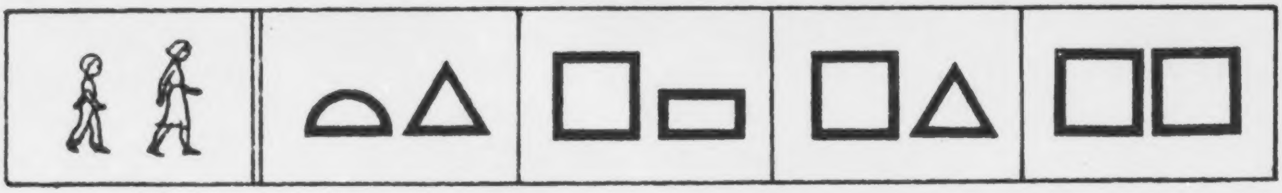


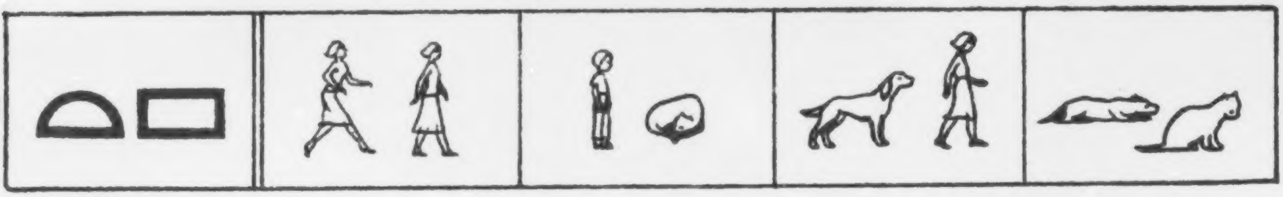


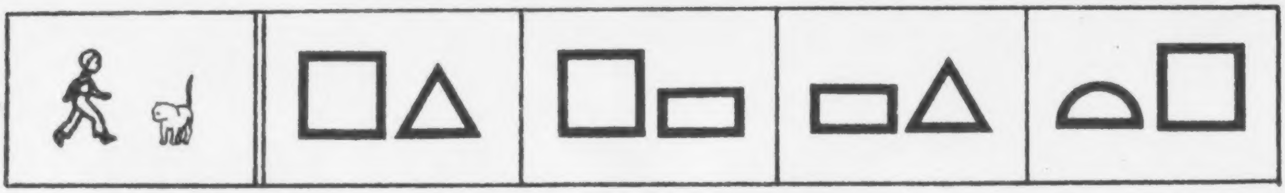


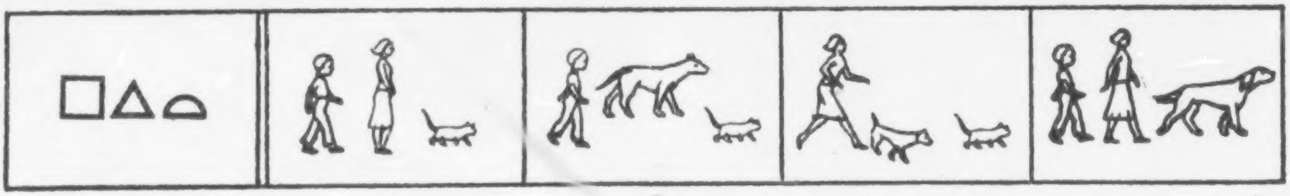


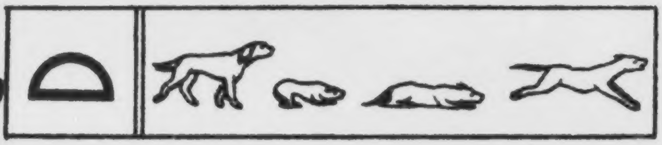
11 

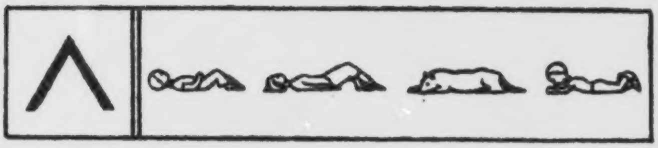
12 

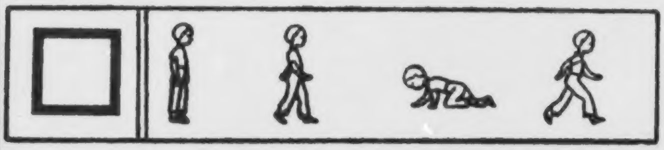
13 

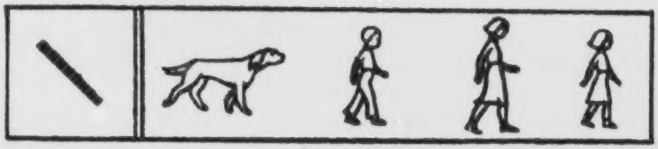
14 

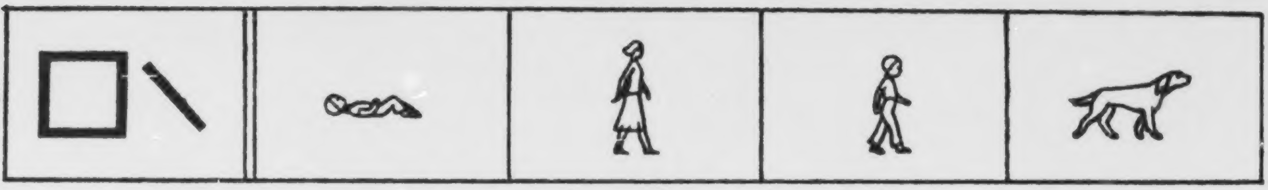
15 

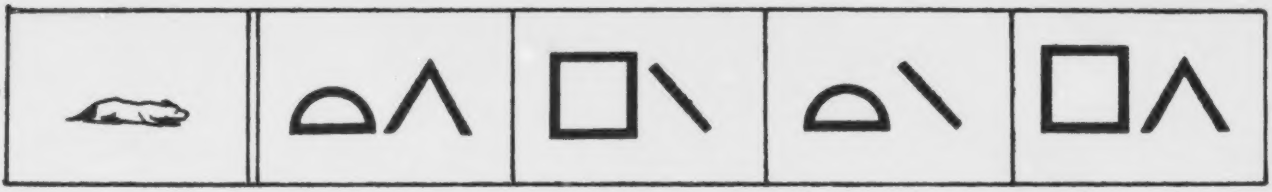


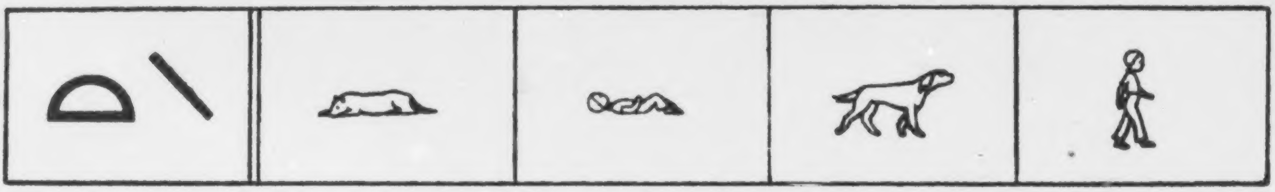


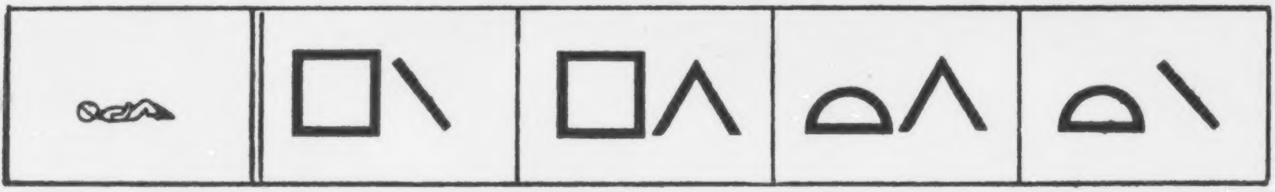


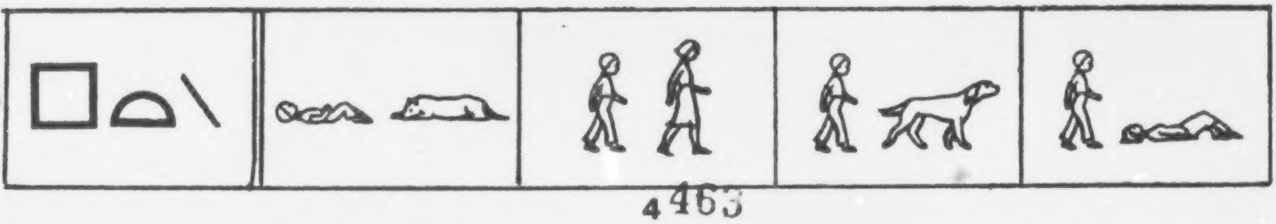


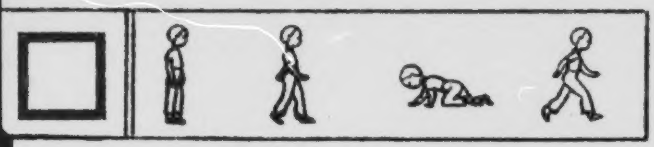
16


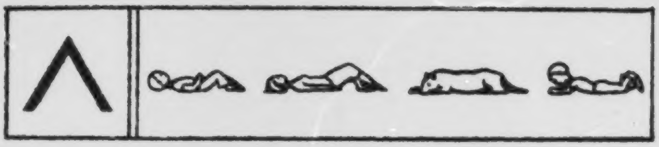
17


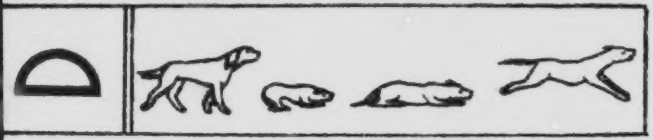
18


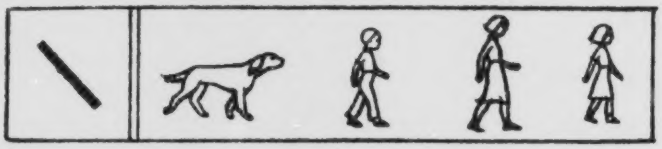
19


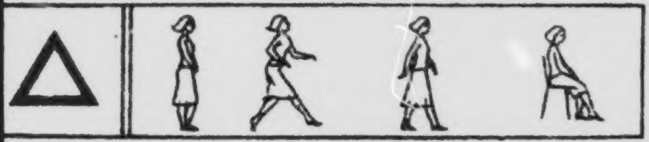
20


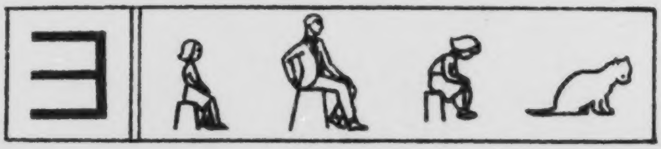


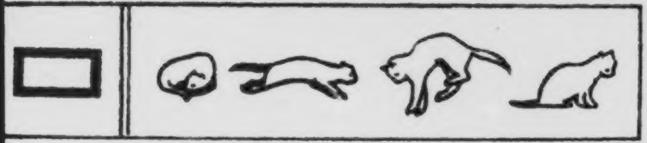


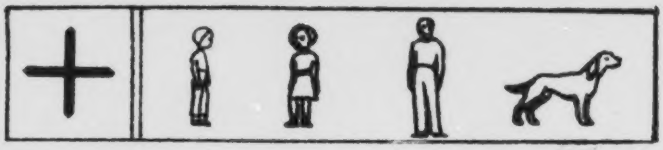


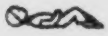
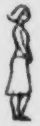
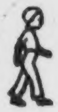
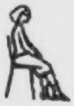
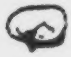
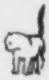
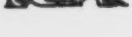
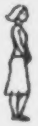
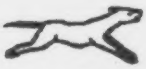
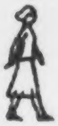
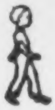
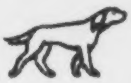


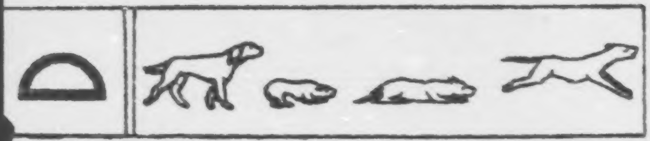


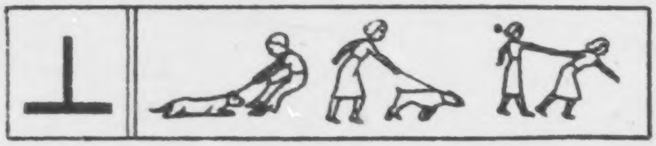


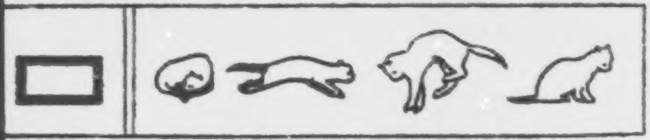


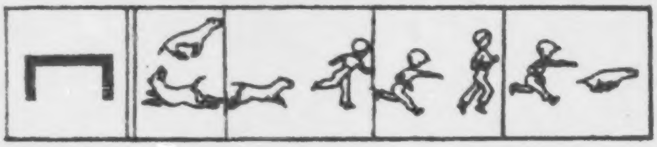


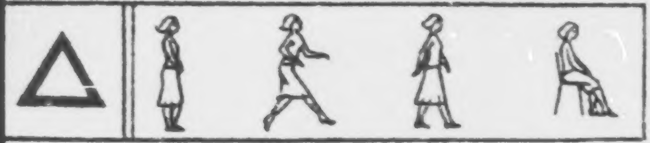
21	□+				
22		△+	□∃	△∃	△\
23	▭∧				
24		△+	□+	△∃	△∧
25	⌓\				

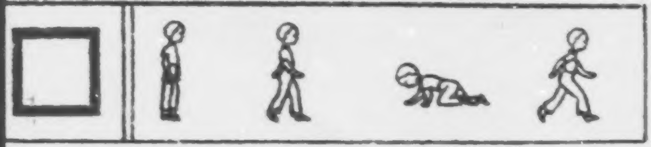


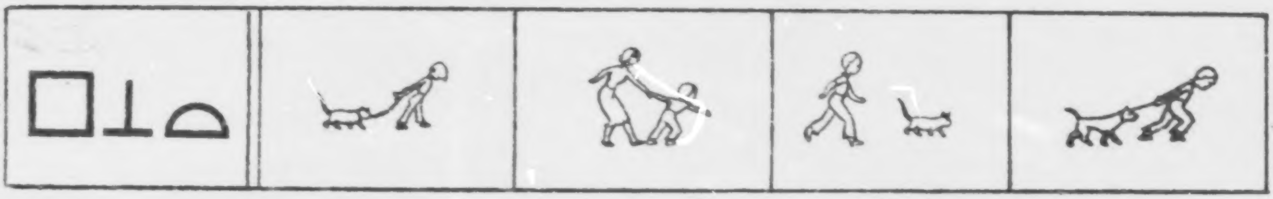


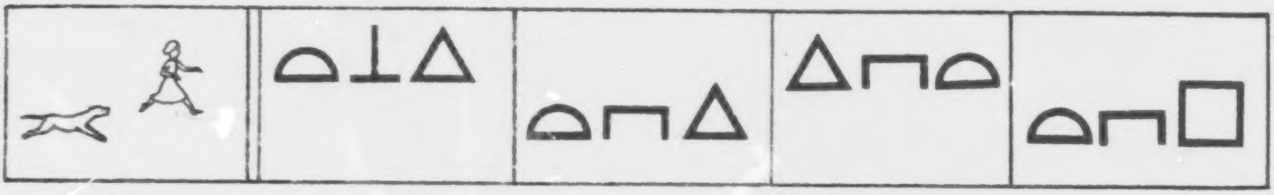


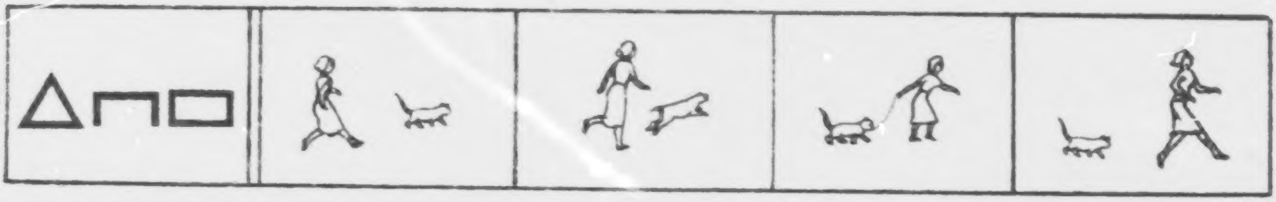


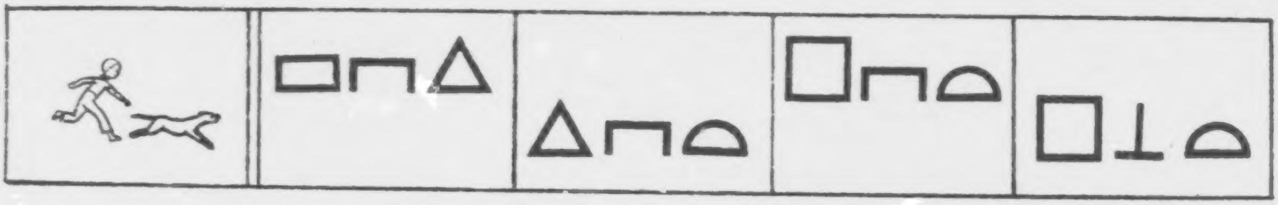


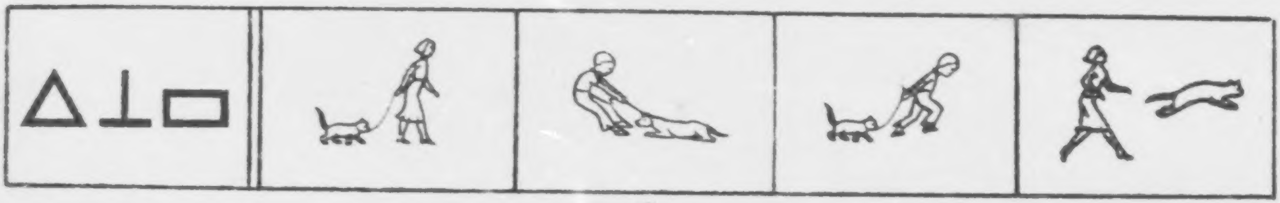


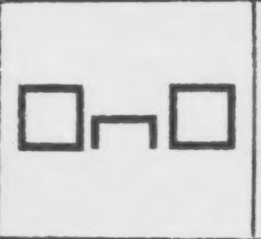
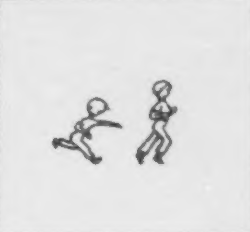
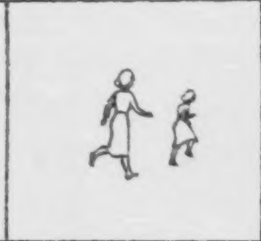
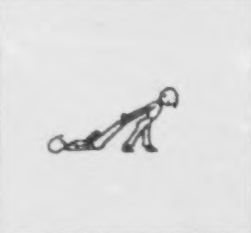
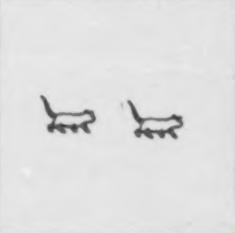
26 

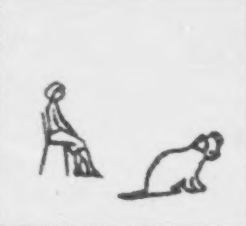
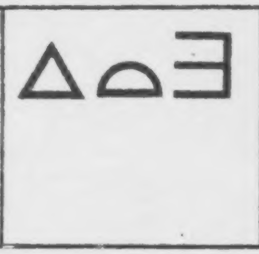
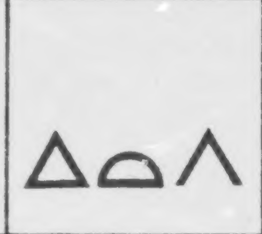
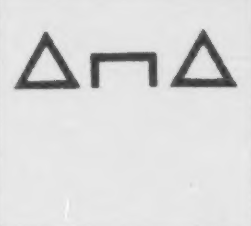
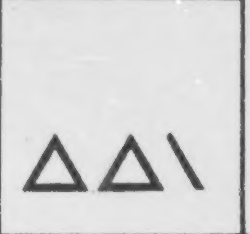
27 

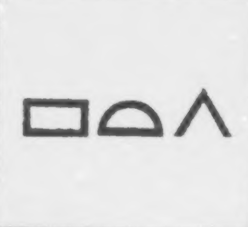
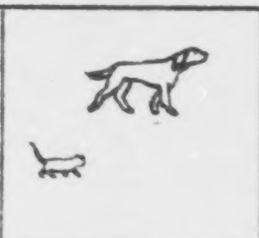
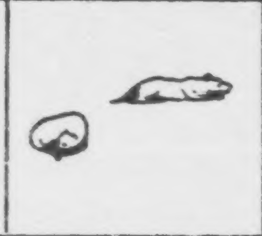
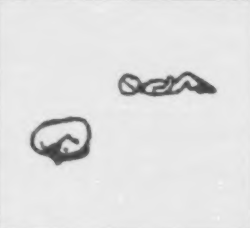
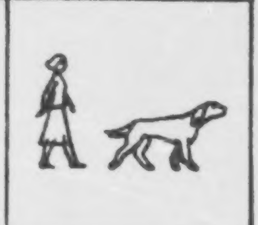
28 

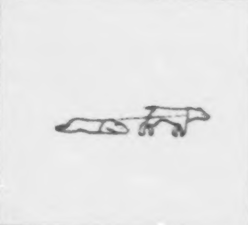
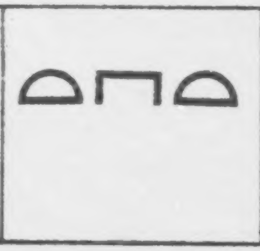
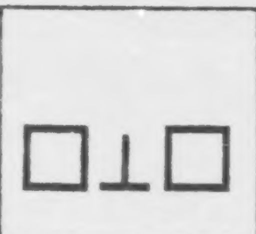
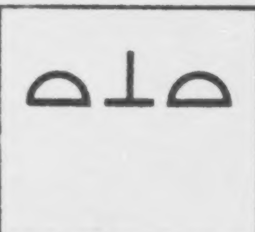
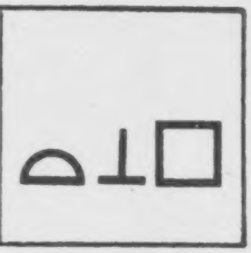
29 

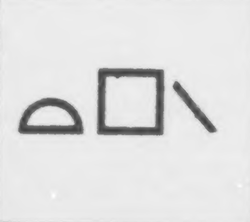
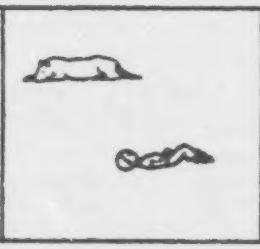
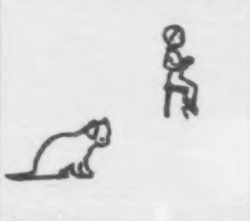
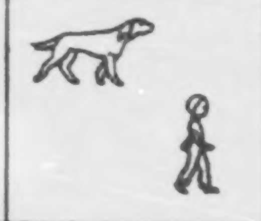
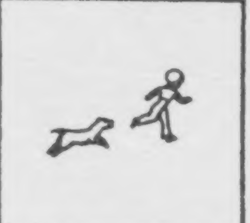
30 

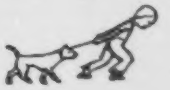
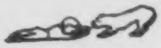
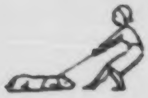
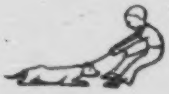
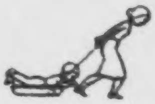
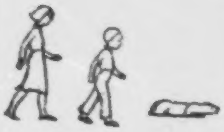
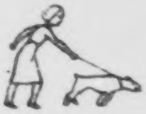
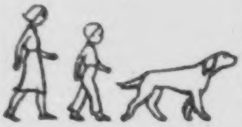
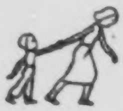
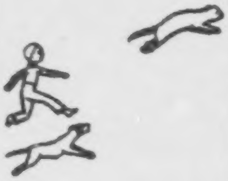
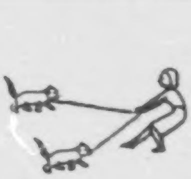
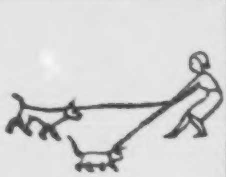
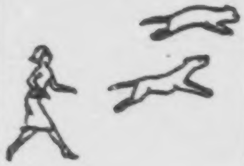
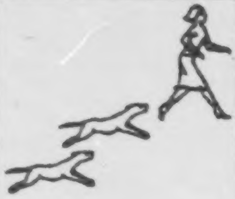
31     

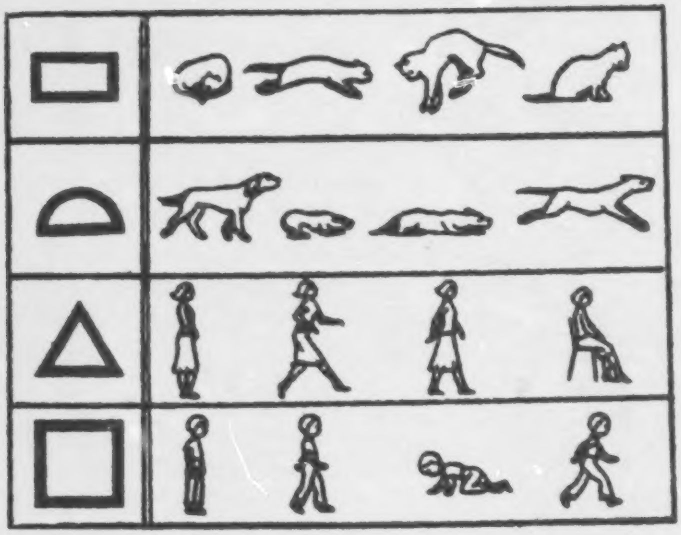
2     

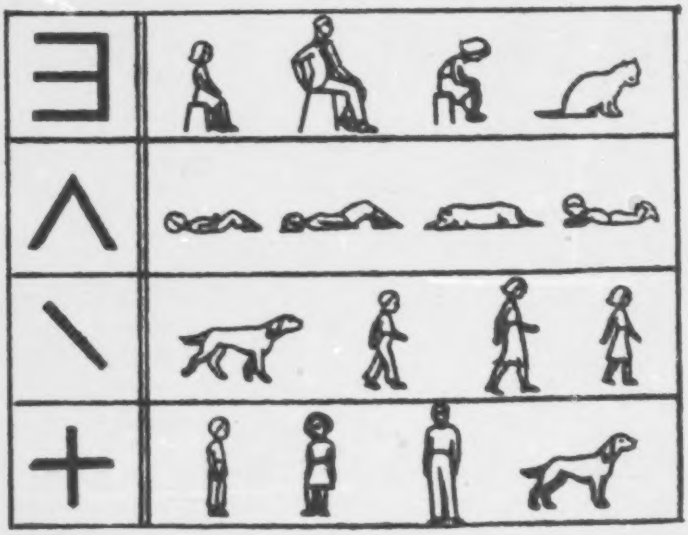
3     

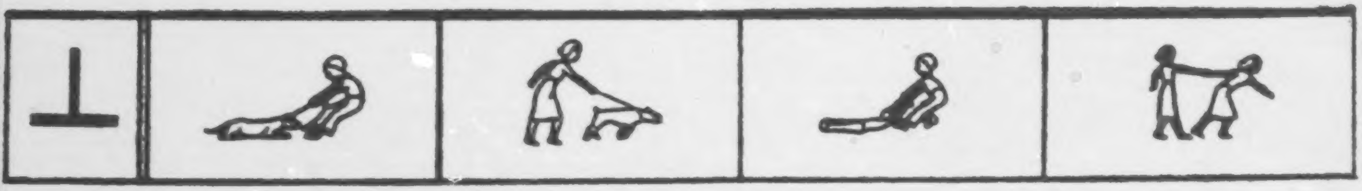
4     

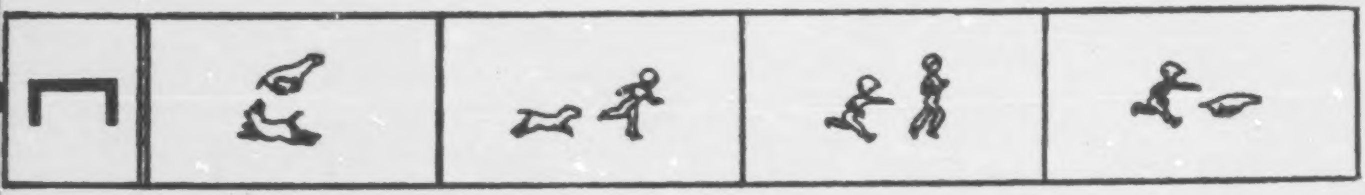
5     

36	□⊥◠∧				
37		△⊥◠∧	□⊥△∧	△⊥□∧	△⊥□\
38	△□◠\				
39		▭▭⊓◠	◠□⊓▭	▭⊓◠⊓	◠□⊥▭
40	△⊥◠▭				









NAME ______________________________ DATE ____________________

GRADE ____________ TEACHER ________________________________

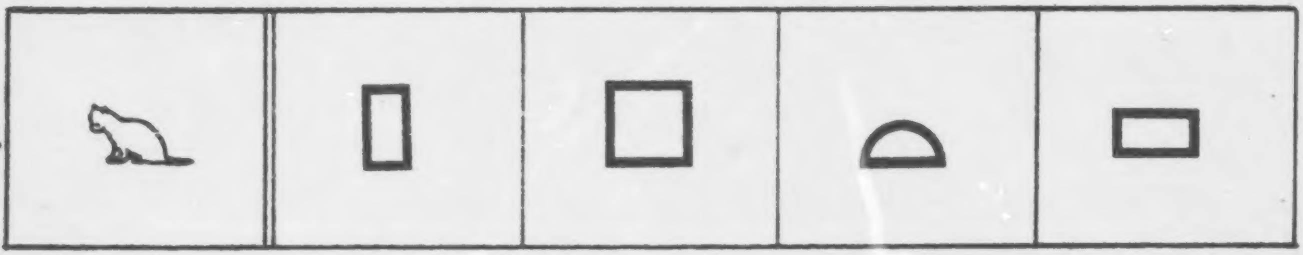
SCHOOL ______________________________

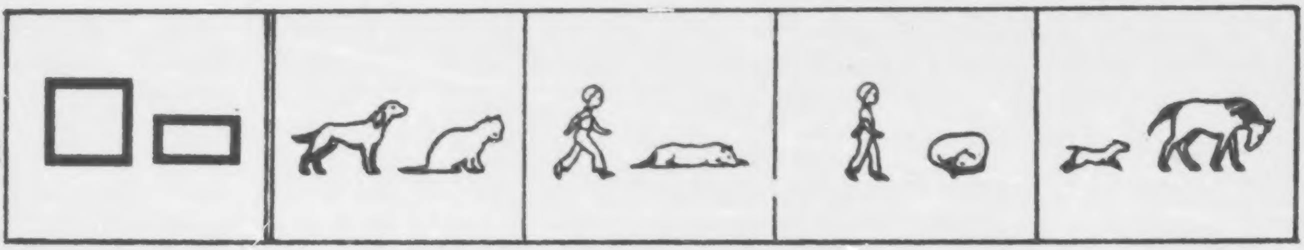
BIRTHDATE (MONTH, DAY, AND YEAR YOU WERE BORN) ______________

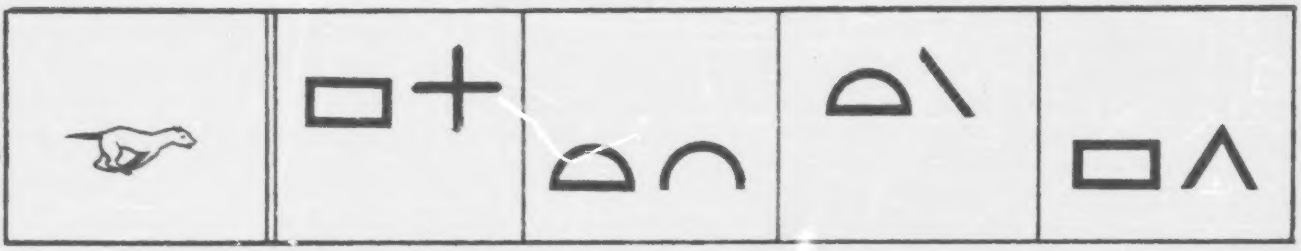
SEX (CHECK ONE) _____ BOY _____ GIRL

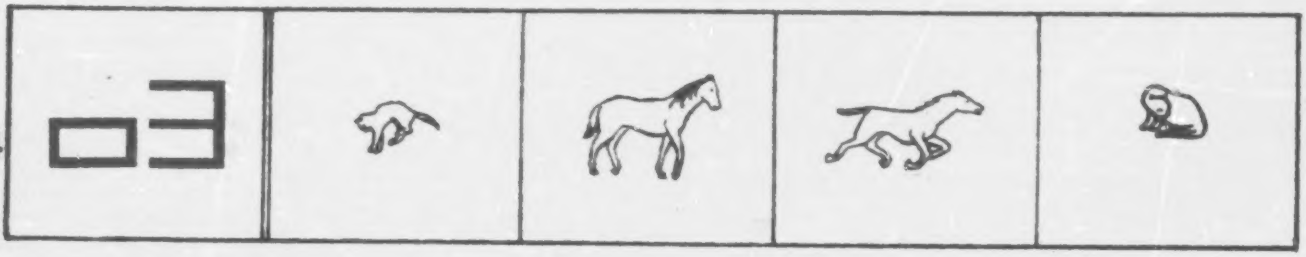
riep

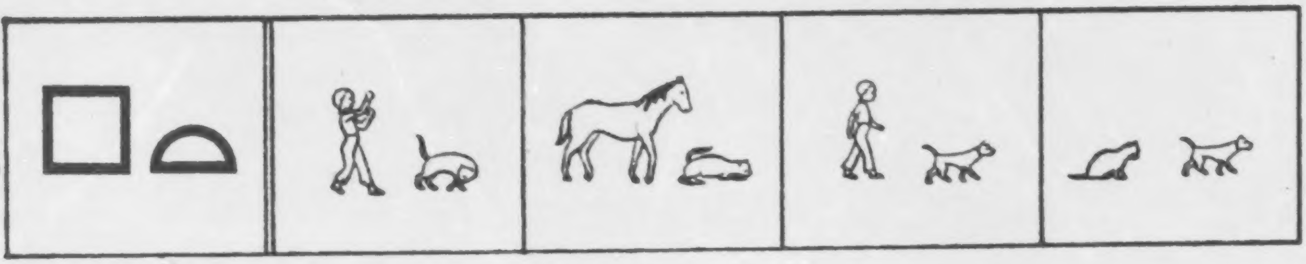
THE PICTURE-WORD GAME

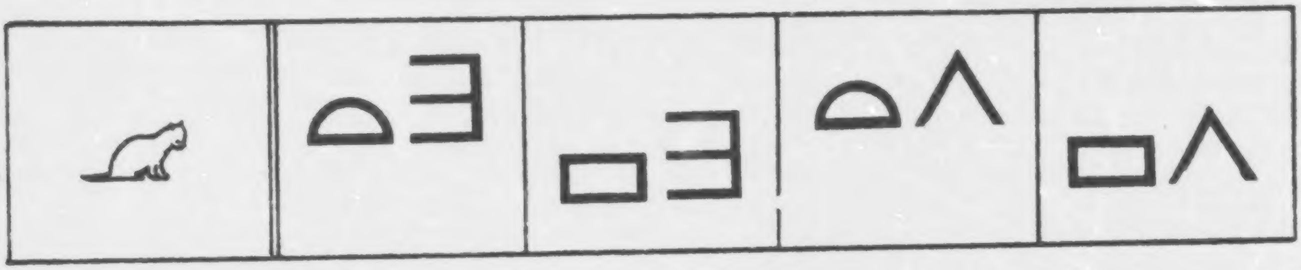
1 

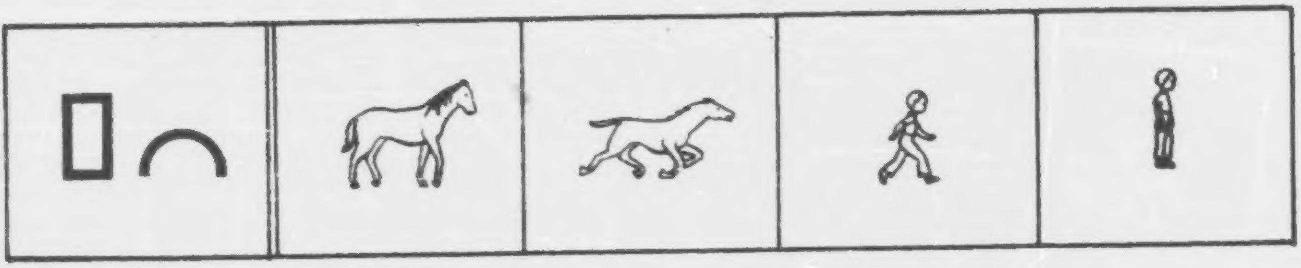
2 

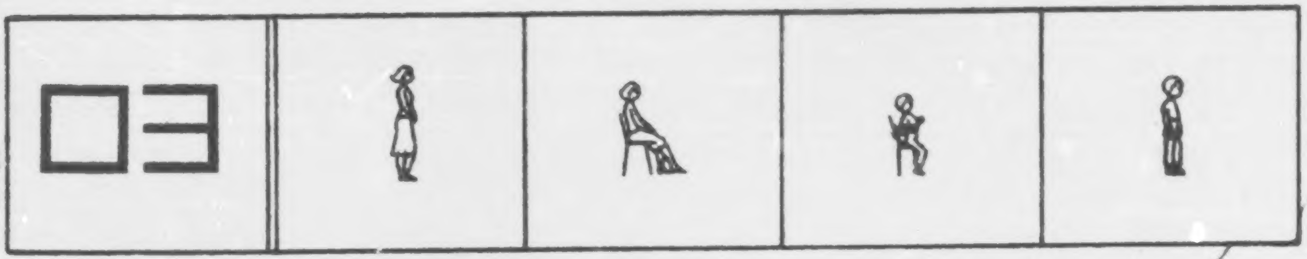
3 

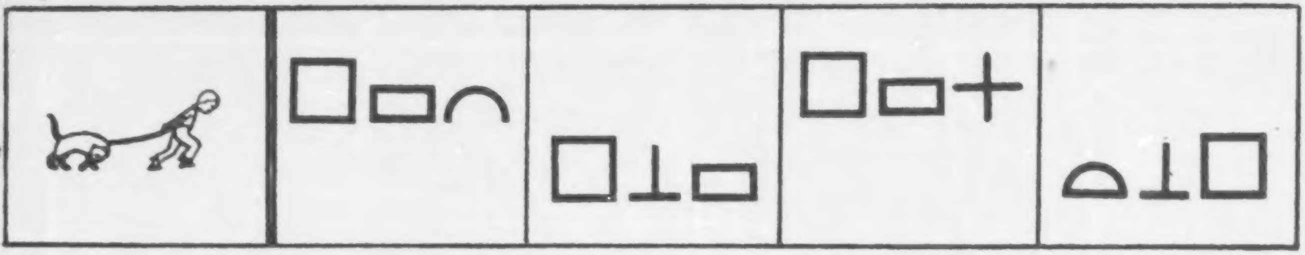
4 

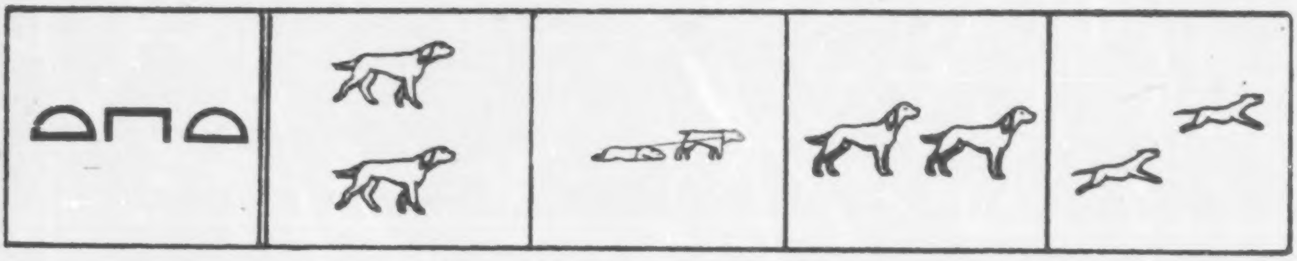
5 

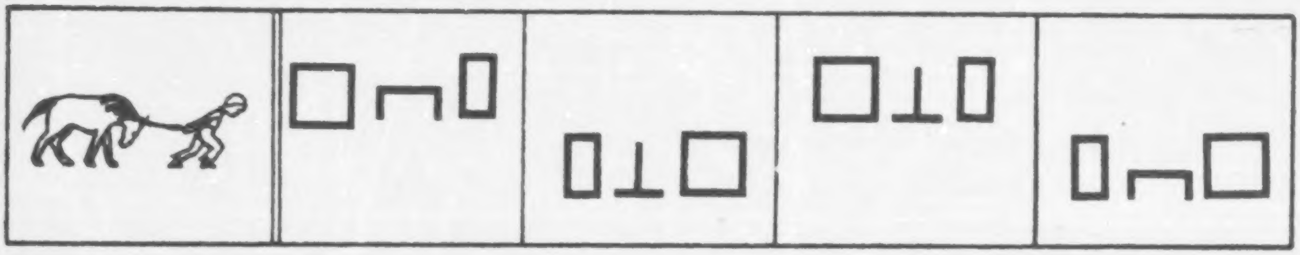
6 

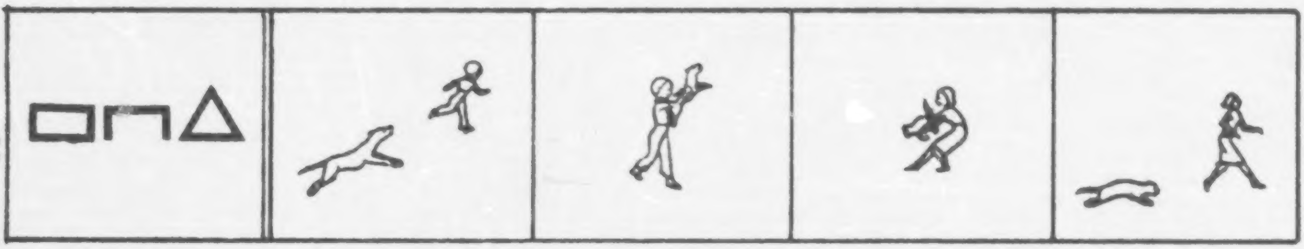
7 

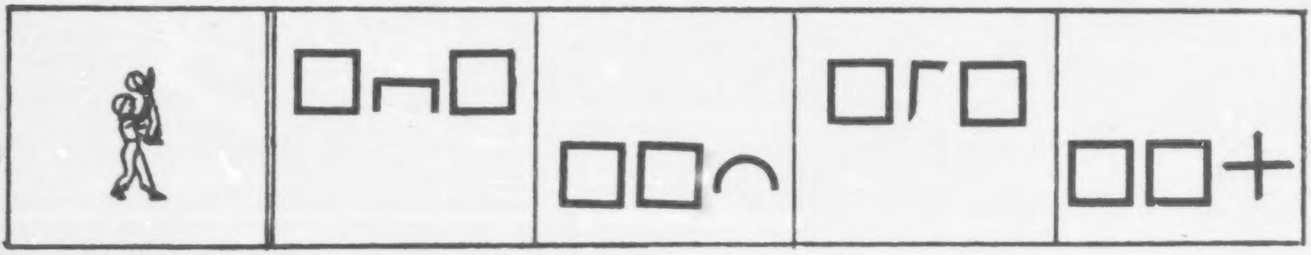
8 

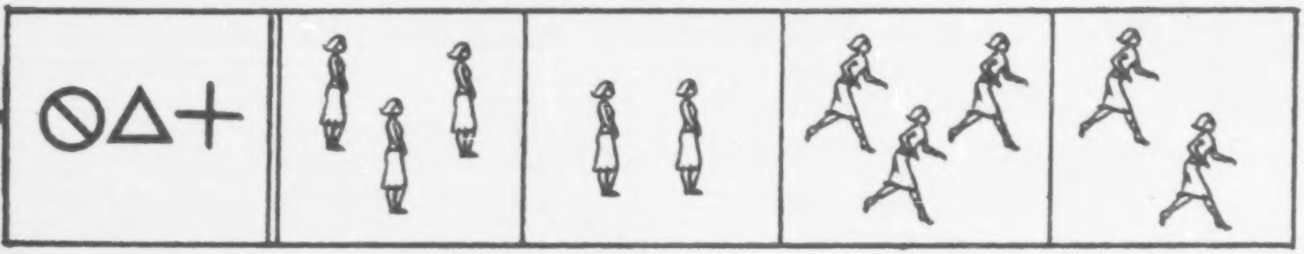




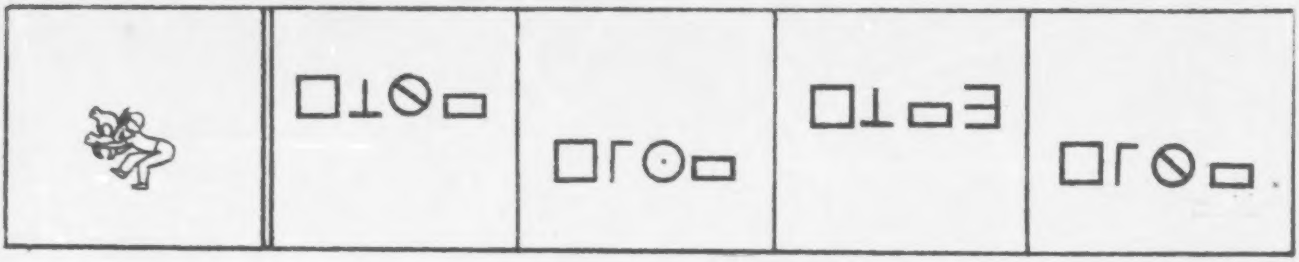




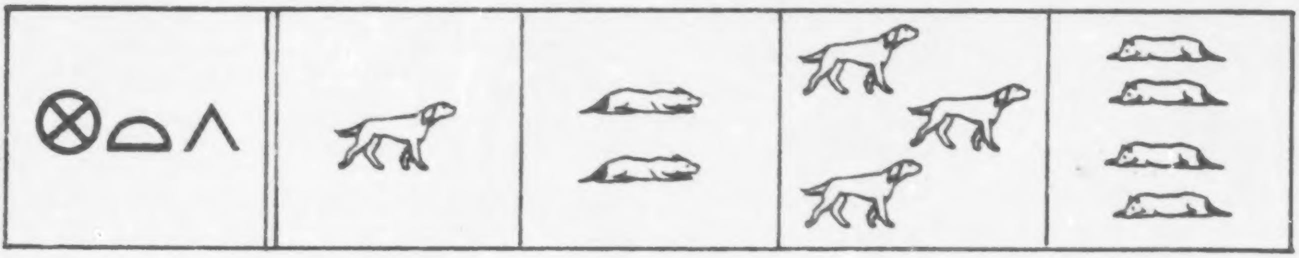




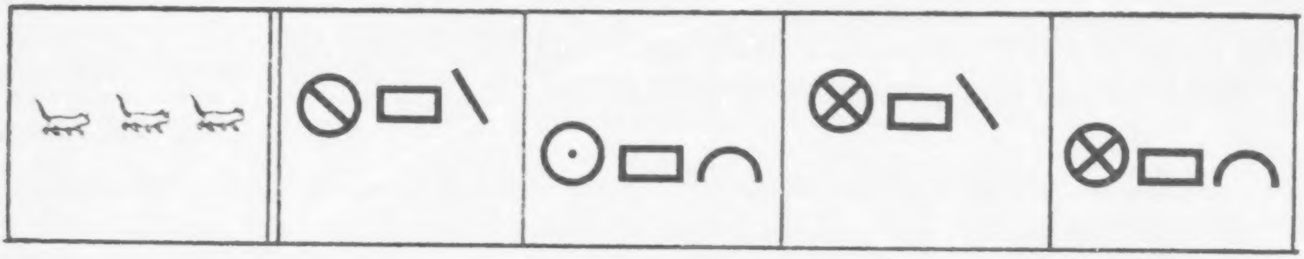
5



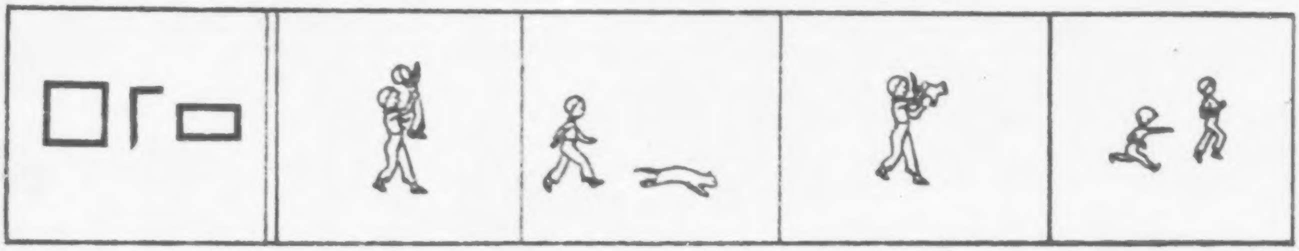
6



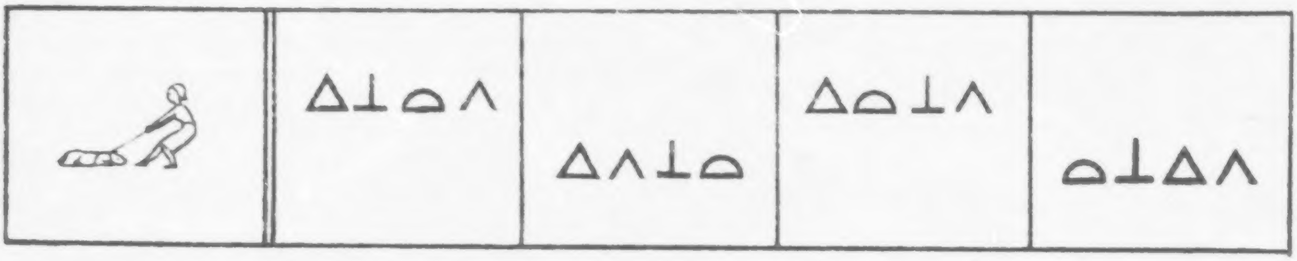
7

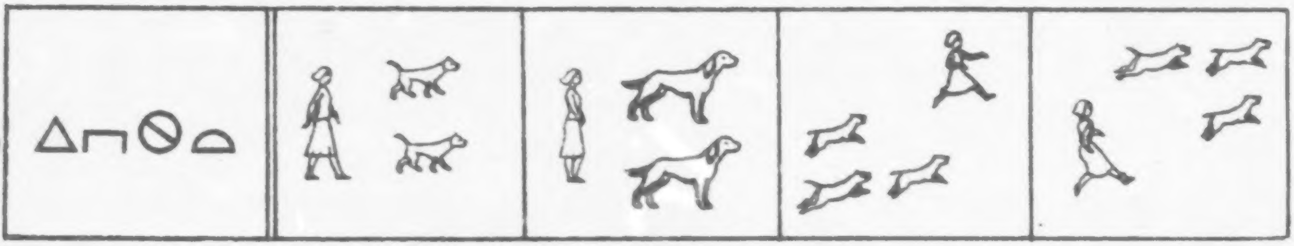


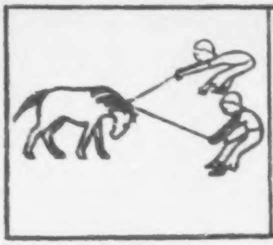
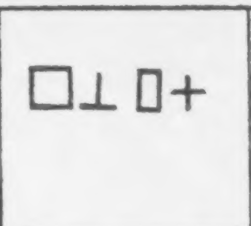
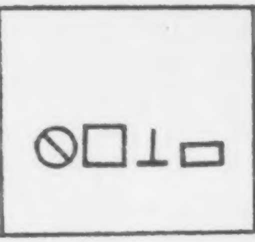
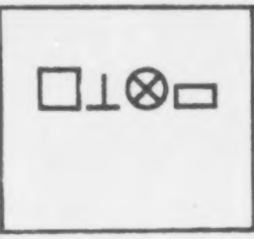
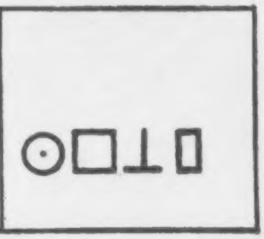
8

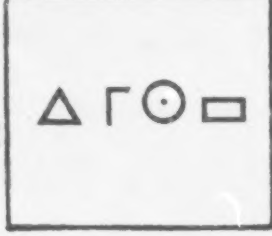
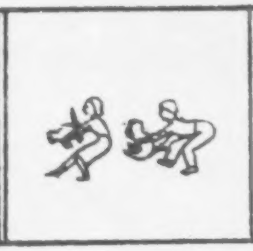
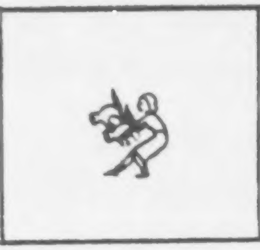
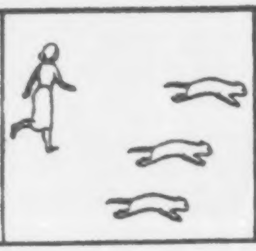
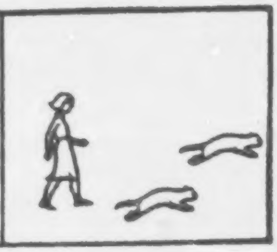


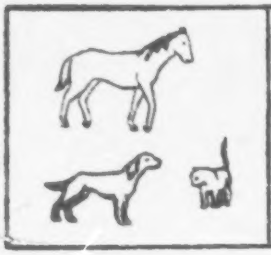
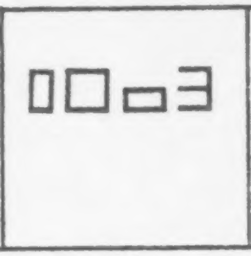
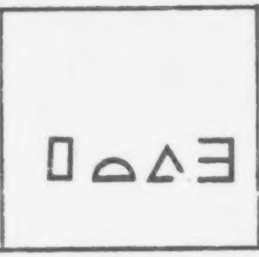
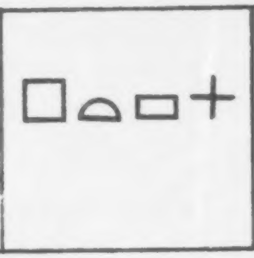
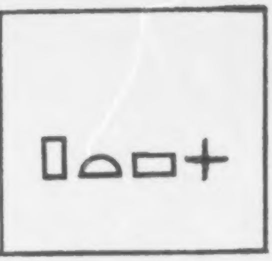
9

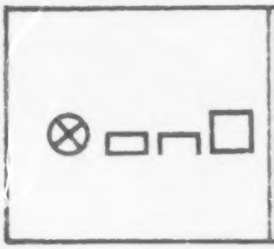
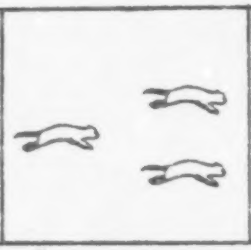
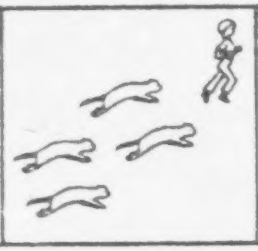
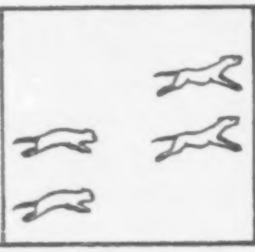
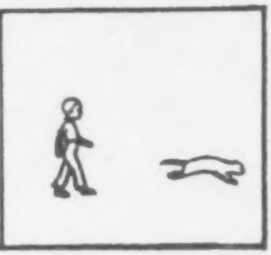


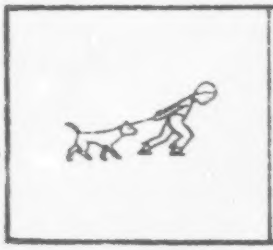
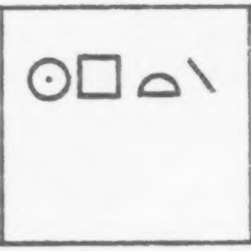
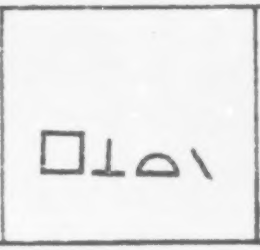
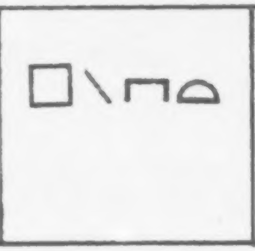
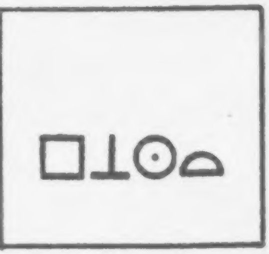


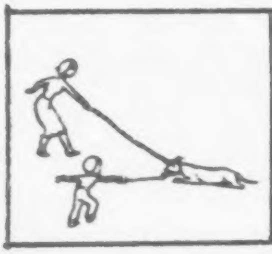
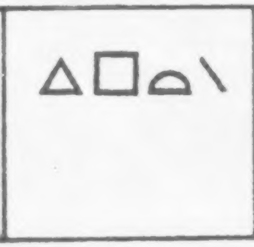
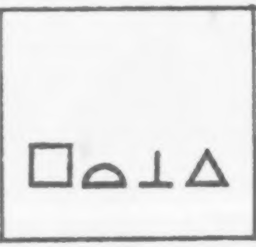
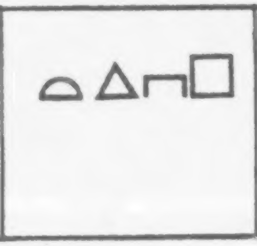
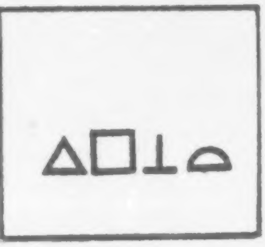
21






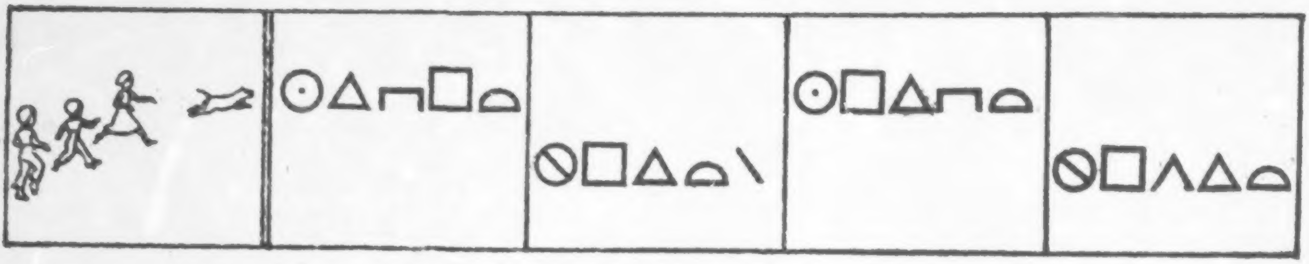
22






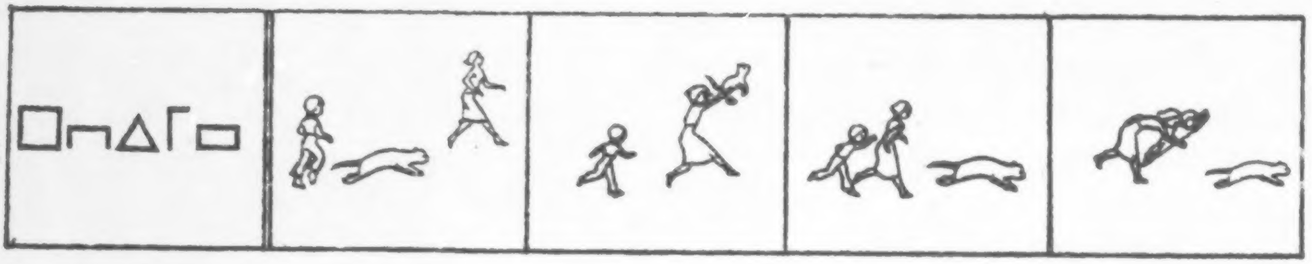
23






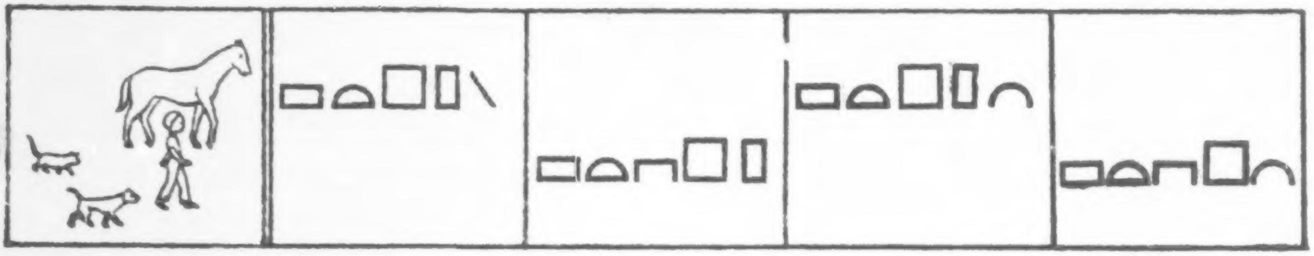
24






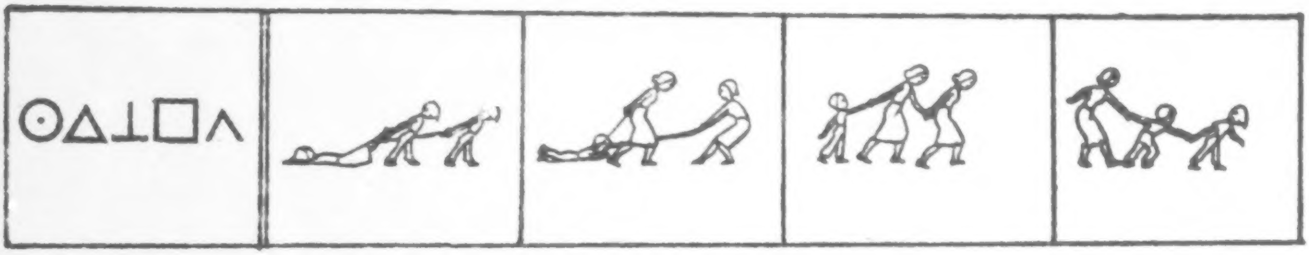
25






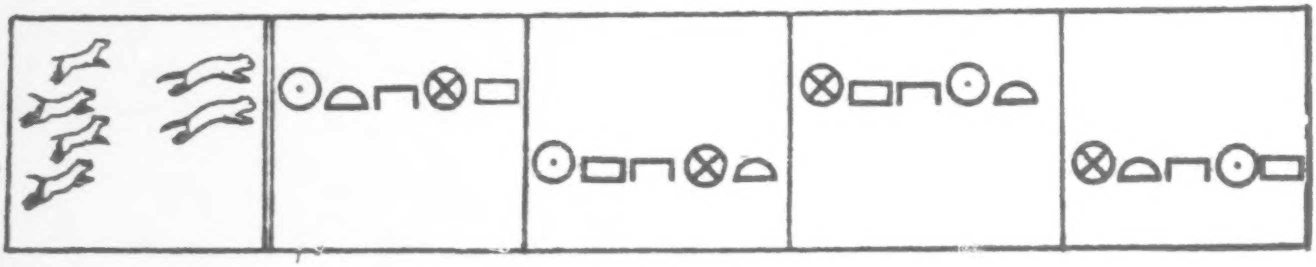
26






27 

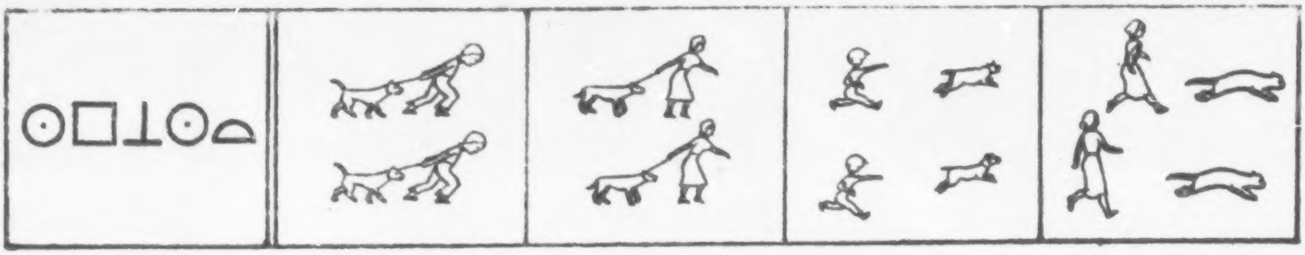
8 

9 

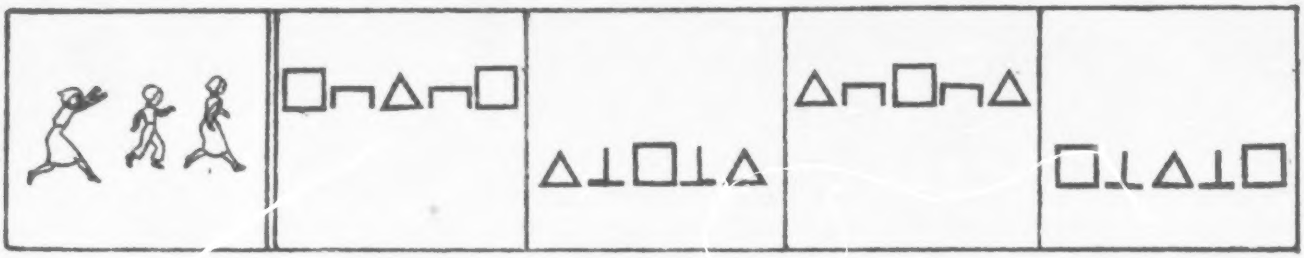
0 

1 

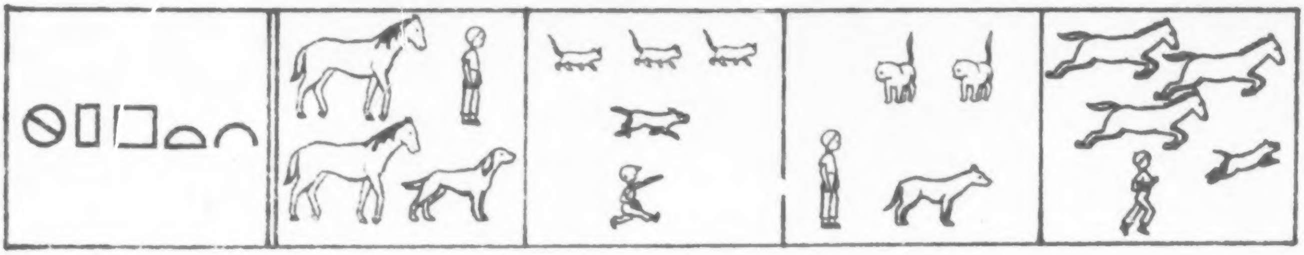
2



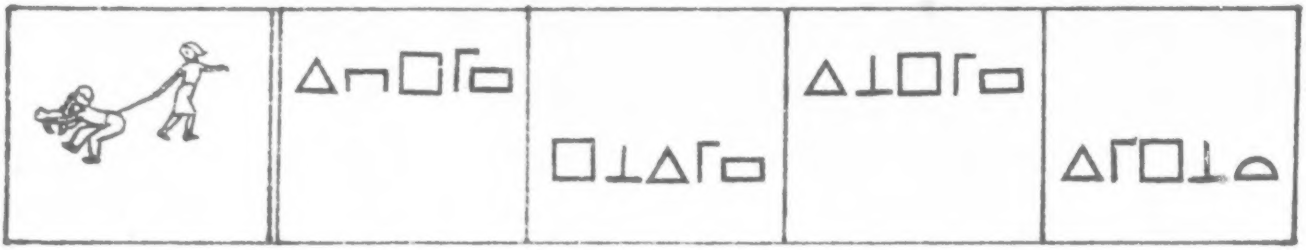
3



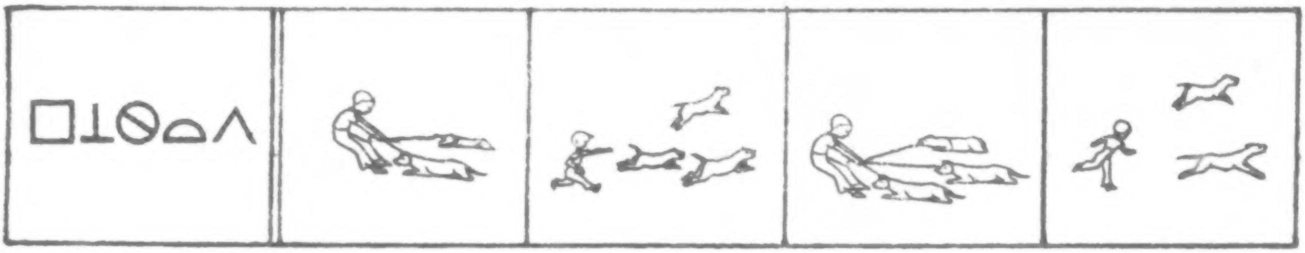
4



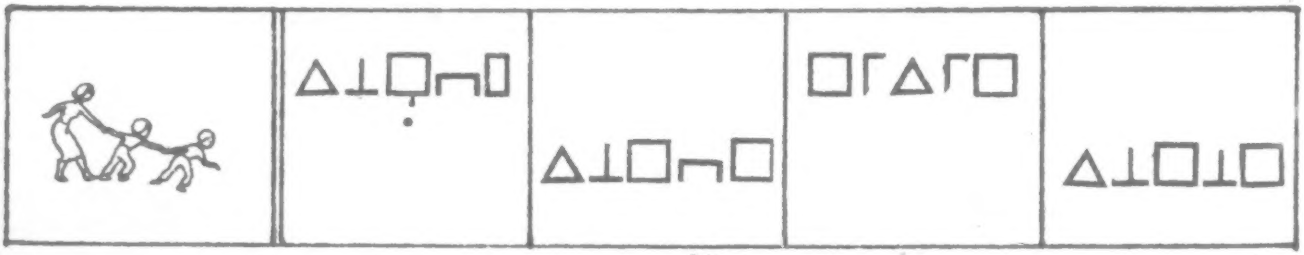
5



6



7



7476